

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

run on: November 21, 2003, 15:19:18 ; Search time 58.8082 Seconds  
(without alignments)  
286.100 Million cell updates/sec

title: US-09-010-377-1

perfect score: 562  
sequence: 1 DIQMTQSPSSLSASVGRVT.....YCLQDNLMTFGQGTKEIK 106

scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1107863 seqs, 158726573 residues

total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database :

```

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	562	100.0	106	AA81321	Humanized anti-VLA
2	562	100.0	106	AAW22412	Humanised alpha-4
3	562	100.0	126	AAW22419	Humanised alpha-4
4	508	90.4	106	AA81328	Mouse anti-VLA-4 a
5	508	90.4	126	AA81326	Mouse VLA-4 antibo
6	508	90.4	126	AA81332	Human VLA-4 reshap
7	508	90.4	126	AAW22409	Alpha-4 integrin m
8	488	85.8	359	AAW22913	Human MCP-3 and mu
9	488	85.8	361	AAW22911	Human IP-10 and mu

10	488	86.8	374	20	AAW22916	Artificial synthe
11	481	85.6	128	15	AA860627	ME1-14 light chain
12	475.5	84.6	234	15	AA813050	CD4-specific CDR-9
13	471.5	83.9	107	16	AA878970	Light chain variab
14	465	82.7	108	17	AA893159	Murine monoclonal
15	460.5	81.9	128	11	AA806252	Variable region of
16	460	81.9	637	13	AA826983	(PRP51)-ETA fusion
17	456	81.1	109	19	AAW26797	Anti-gp54 MAb T16
18	455	81.0	240	16	AA85495	SCF(FWP51). Sync
19	455	81.0	241	20	AAW21882	Amino acid sequenc
20	455	81.0	245	19	AAW26800	Anti-gp54 MAb T16
21	453.5	80.7	107	16	AA81322	Humanized VLA-4 an
22	453.5	80.7	107	18	AAW22422	Humanised alpha-4
23	453	80.6	109	12	AA813558	Murine OKT4A light
24	452	80.4	241	13	AA826981	FWP51 fusion prote
25	451.5	80.3	108	16	AA85163	Human REI monoclon
26	451.5	80.3	129	15	AA847207	Human/murine IL-1
27	451	80.2	355	18	AAW35133	R. pipiens recombi
28	449.5	80.0	107	19	AAW86805	Variable light dom
29	449.5	80.0	107	19	AAW70825	Humanised murine a
30	449.5	80.0	107	23	ABP61194	Humanised anti-VEG
31	446.5	79.4	108	18	AAW10231	CDR-grafted light
32	446.5	79.4	110	19	AAW70673	Anti-VEGF humanise
33	446.5	79.4	110	23	ABP61242	Humanised anti-VEG
34	446.5	79.4	234	18	AAW10233	TF8-5G9 CDR-grafte
35	446.5	79.4	237	19	AAW70703	Protein encoded by
36	446.5	79.4	650	23	ABP61341	Phage-display anti
37	445.5	79.3	107	19	AAW86804	Variable light dom
38	445.5	79.3	107	19	AAW70623	Humanised murine a
39	445.5	79.3	107	23	ABP61192	Humanised anti-VEG
40	444.5	79.1	107	20	AAW87455	Humanised anti-alp
41	444.5	79.1	107	22	AA862087	Human VI consensus
42	444.5	79.1	107	22	AA860400	Consensus human li
43	444.5	79.1	107	22	AA861585	Human variable lig
44	444.5	79.1	108	19	AAW70622	Human consensus fr
45	444.5	79.1	108	21	AAW82345	Human consensus se

#### ALIGNMENTS

RESULT 1  
AA81321  
ID AA81321 standard; Protein; 106 AA.  
XX AC AA81321;  
XX DT 02-APR-1996 (first entry)  
XX DE Humanized anti-VLA-4 antibody 21.6 light chain variable region, La.  
XX KW Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;  
XX OS antibody engineering.  
XX OS Chimeric Mus musculus.  
XX OS Chimeric Homo sapiens.  
XX PN WO9519790-A1.  
XX PD 27-JUL-1995.  
XX PF 25-JAN-1995; 95WO-US01219.  
XX PR 25-JAN-1994; 94US-0186269.  
XX (ATHE-) ATHENA NEUROSCIENCES INC.  
XX Bendi MM, Jones TS, Leger OJ, Saldanha J;  
XX WPI; 1995-269276/35.  
XX New humanised antibodies against VLA-4 - used for inhibiting  
XX leukocyte adhesion to endothelial cells, partic. for treating



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

on: November 21, 2003, 15:21:54 ; Search time 19.9658 Seconds  
(without alignments)  
224.632 Million cell updates/sec

le: US-09-010-377-1

fect score: 562

uence: 1 DIQWTQSPSLASVGRVT.....YCLQYDNLWTFQGTKEIK 106

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

rchd: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

imum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase :

Issued Patents AA:\*\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	562	100.0	106	2	US-08-561-521-7	Sequence 7, Appl
2	562	100.0	106	5	PCT-US95-01219-7	Sequence 7, Appl
3	508	90.4	106	2	US-08-561-521-5	Sequence 5, Appl
4	508	90.4	106	5	PCT-US95-01219-5	Sequence 5, Appl
5	508	90.4	126	2	US-08-561-521-2	Sequence 2, Appl
6	508	90.4	126	2	US-08-561-521-15	Sequence 15, Appl
7	508	90.4	126	5	PCT-US95-01219-2	Sequence 2, Appl
8	508	90.4	126	5	PCT-US95-01219-15	Sequence 15, Appl
9	488	86.8	359	4	US-09-646-028-16	Sequence 16, Appl
10	488	86.8	361	4	US-09-646-028-13	Sequence 13, Appl
11	488	86.8	374	4	US-09-646-028-15	Sequence 15, Appl
12	481	85.6	128	1	US-08-339-582-4	Sequence 4, Appl
13	471.5	83.9	107	2	US-08-898-366-22	Sequence 22, Appl
14	460	81.9	637	1	US-08-235-938-16	Sequence 16, Appl
15	460	81.9	637	2	US-08-465-473B-16	Sequence 16, Appl
16	455	81.0	241	1	US-08-235-838-11	Sequence 11, Appl
17	455	81.0	241	1	US-08-465-473B-11	Sequence 11, Appl
18	454.5	80.9	108	2	US-08-602-725-29	Sequence 29, Appl
19	453.5	80.7	107	2	US-08-561-521-8	Sequence 8, Appl
20	453.5	80.7	107	5	PCT-US95-01219-8	Sequence 8, Appl
21	451	80.2	355	3	US-08-875-811-57	Sequence 57, Appl
22	448.5	79.8	107	2	US-07-934-373C-17	Sequence 17, Appl
23	448.5	79.8	107	3	US-08-437-642B-17	Sequence 17, Appl
24	448.5	79.8	107	4	US-08-146-206C-17	Sequence 17, Appl
25	448.5	79.8	107	5	PCT-US93-07832-17	Sequence 17, Appl
26	444.5	79.1	107	2	US-07-934-373C-18	Sequence 18, Appl
27	444.5	79.1	107	3	US-08-437-642B-18	Sequence 18, Appl

28 444.5 79.1 107 4 US-08-146-206C-18 Sequence 18, Appl  
29 444.5 79.1 107 5 PCT-US93-07832-18 Sequence 18, Appl  
30 444.5 79.1 108 3 US-08-974-893-3 Sequence 3, Appl  
31 444.5 79.1 111 1 US-08-137-117D-67 Sequence 67, Appl  
32 444.5 79.1 111 2 US-08-436-717-67 Sequence 67, Appl  
33 444.5 79.1 126 1 US-08-137-117D-71 Sequence 71, Appl  
34 444.5 79.1 126 2 US-08-436-717-71 Sequence 71, Appl  
35 443.5 78.9 107 2 US-08-561-521-6 Sequence 6, Appl  
36 443.5 78.9 107 2 US-08-652-558-34 Sequence 34, Appl  
37 443.5 78.9 107 4 US-09-025-203-15 Sequence 15, Appl  
38 443.5 78.9 107 5 PCT-US95-01219-6 Sequence 6, Appl  
39 443.5 78.9 108 2 US-08-070-116A-7 Sequence 7, Appl  
40 443.5 78.9 108 2 US-08-116-247-9 Sequence 9, Appl  
41 443.5 78.9 108 4 US-08-557-050-7 Sequence 7, Appl  
42 441.5 78.6 107 2 US-08-318-157B-6 Sequence 6, Appl  
43 440.5 78.4 107 2 US-08-652-558-2 Sequence 2, Appl  
44 440.5 78.4 109 2 US-07-934-373C-3 Sequence 3, Appl  
45 440.5 78.4 109 3 US-08-437-642B-3 Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-08-561-521-7  
; Sequence 7, Application US/08561521  
; Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; TITLE OF INVENTION: Adhesion Molecule VLA-4  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-561-521-7

Query Match 100.0%; Score 562; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.6e-45;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60  
1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60  
61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106  
61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106

RESULT 2

CT-US95-01219-7

Sequence 7, Application PC/TUS9501219

GENERAL INFORMATION:

APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01219  
FILING DATE: 25-JAN-1995  
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/186,269  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.

REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

CT-US95-01219-7

Query Match 100.0%; Score 562; DB 5; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.6e-45;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60  
1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60  
61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106  
61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106

RESULT 3

S-08-561-521-5

Sequence 5, Application US/08561521

Patent No. 5840299

GENERAL INFORMATION:

APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/561,521  
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/186,269A  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.

REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 15270-14

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-543-9600

TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-561-521-5

Query Match 90.4%; Score 508; DB 2; Length 106;  
Best Local Similarity 88.7%; Pred. No. 1.7e-40;  
Matches 94; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60  
1 DIQMTQSPSSLSASVGDRTVITCKTSQDINKYMAWYQQTGKAPRLLIHYTSALQPGIPS 60

61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106

61 RFGSGSGRDYFTTSSLPQEDATYYCLOYDNLWTFGGQTKVEIK 106

RESULT 4

PCT-US95-01219-5

Sequence 5, Application PC/TUS9501219

GENERAL INFORMATION:

APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: November 21, 2003, 15:27:54 ; Search time 39.5685 Seconds  
(without alignments)  
489.058 Million cell updates/sec

Title: US-09-010-377-1

Perfect score: 562

Sequence: 1 DIQWTSQSSLSASVGRVT.....YCLQYDNLWTFGQTKVEIK 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	86.8	359	12	US-10-335-394-16
2	488	86.8	361	12	US-10-335-394-13
3	488	86.8	374	12	US-10-335-394-15
4	479.5	85.3	108	9	US-09-229-200A-11
5	479.5	85.3	108	9	US-09-229-200A-15
6	466.5	83.0	108	9	US-09-229-200A-16
7	449.5	80.0	107	9	US-09-056-160B-15
8	449.5	80.0	107	12	US-10-234-671-15
9	446.5	79.4	110	9	US-09-056-160B-103
10	446.5	79.4	110	12	US-10-234-671-101
11	446.5	79.4	237	9	US-09-056-160B-100
12	446.5	79.4	237	12	US-10-234-671-100
13	446.5	79.4	491	14	US-10-011-125-2
14	445.5	79.3	107	9	US-09-056-160B-13
15	445.5	79.3	107	12	US-10-234-671-13

16	444.5	79.1	107	12	US-10-223-880-15	Sequence 15, Appl
17	444.5	79.1	107	15	US-10-268-501-5	Sequence 5, Appl
18	444.5	79.1	108	9	US-09-056-160B-12	Sequence 12, Appl
19	444.5	79.1	108	12	US-10-234-671-12	Sequence 12, Appl
20	444.5	79.1	108	12	US-09-795-798-3	Sequence 3, Appl
21	444.5	79.1	109	9	US-09-811-123-6	Sequence 6, Appl
22	444.5	79.1	110	12	US-10-044-896-4	Sequence 4, Appl
23	443.5	78.9	107	10	US-09-999-025-15	Sequence 15, Appl
24	443.5	78.9	107	10	US-09-999-040-15	Sequence 15, Appl
25	443.5	78.9	107	11	US-09-998-817-15	Sequence 15, Appl
26	443.5	78.9	107	11	US-09-999-021-15	Sequence 15, Appl
27	443.5	78.9	107	15	US-10-040-997-15	Sequence 15, Appl
28	443.5	78.9	108	9	US-09-229-200A-14	Sequence 14, Appl
29	443.5	78.9	108	12	US-10-428-662-100	Sequence 100, App
30	443.5	78.9	108	12	US-10-422-049-7	Sequence 7, Appl
31	443.5	78.9	108	15	US-10-267-286A-7	Sequence 7, Appl
32	443.5	78.9	110	9	US-09-056-160B-105	Sequence 105, App
33	443.5	78.9	110	12	US-10-234-671-103	Sequence 103, App
34	443	78.8	109	9	US-09-229-200A-7	Sequence 7, Appl
35	442.5	78.7	108	9	US-09-056-160B-8	Sequence 8, Appl
36	442.5	78.7	108	14	US-10-153-159-2	Sequence 2, Appl
37	442.5	78.7	108	14	US-10-153-159-16	Sequence 16, Appl
38	442.5	78.7	108	15	US-10-153-176-2	Sequence 2, Appl
39	442.5	78.7	108	15	US-10-153-176-16	Sequence 16, Appl
40	442.5	78.7	110	12	US-10-234-671-8	Sequence 8, Appl
41	442.5	78.7	131	15	US-10-207-655-256	Sequence 256, App
42	442.5	78.7	266	15	US-10-207-655-260	Sequence 260, App
43	442.5	78.7	550	15	US-10-207-655-270	Sequence 270, App
44	441.5	78.6	107	9	US-09-253-794-6	Sequence 6, Appl
45	439.5	78.2	108	14	US-10-153-159-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-10-335-394-16  
; Sequence 16, Application US/10335394  
; Publication No. US20030138452A1  
; GENERAL INFORMATION:  
; APPLICANT: Kwak, Larry  
; APPLICANT: Bitagyn, Arya  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF  
; FILE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES  
; FILE REFERENCE: 14014.0316/P  
; CURRENT APPLICATION NUMBER: US/10/335,394  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/646,028  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 60/077,745  
; PRIOR FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct  
US-10-335-394-16

Query Match 86.8%; Score 488; DB 12; Length 359;  
Best Local Similarity 84.9%; Pred. No. 6.8e-37;  
Matches 90; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
QY 1 DIQWTSQSSLSASVGRVTCTCKTSQDINKYMWYQQTGKAPRLLIHYTSLALOPGIPS 60  
DB 233 DIQWTSQSSLSASVGRVTCTCKTSQDINKYMWYQQTGKAPRLLIHYTSLALOPGIPS 292  
QY 61 RFSGSGSGRDYFTFTISSLPEDIATYYCIQYDNLWTFGQTKVEIK 106  
DB 293 RFSGSGSGRDYFTFTISSLPEDIATYYCIQYDNLWTFGQTKVEIK 338

RESULT 2  
US-10-335-394-13  
; Sequence 13, Application US/10335394  
; Publication No. US20030138452A1  
; GENERAL INFORMATION:  
; APPLICANT: Kwak, Larry  
; APPLICANT: Biragyn, Arya  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF  
; TITLE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES  
; FILE REFERENCE: 14014.0316/P  
; CURRENT APPLICATION NUMBER: US/10/335,394  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/646,028  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 60/077,745  
; PRIOR FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct  
US-10-335-394-13  
Query Match 86.8%; Score 488; DB 12; Length 361;  
Best Local Similarity 84.9%; Pred. No. 6.8e-37;  
Matches 90; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 DIQMTQSPSSLSASVGRVITCKTSQDINKYMWYQQTGKAPRLLIHYTSLALQGPIS 60  
Db 235 DIQMTQSPSSLSASLGKVIITCKASQDINKYIAWYQHKGPRLLIHYTSLALQGPIS 294  
Qy 61 RFSGSGGRDYTTFTISSLPEDIATYYCLOYDNLWTFGGQTKVEIK 106  
Db 295 RFSGSGGRDYTSFISNLEPEDIATYYCLOYDNLWTFGGGKLEIK 340  
RESULT 3  
US-10-335-394-15  
; Sequence 15, Application US/10335394  
; Publication No. US20030138452A1  
; GENERAL INFORMATION:  
; APPLICANT: Kwak, Larry  
; APPLICANT: Biragyn, Arya  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF  
; TITLE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES  
; FILE REFERENCE: 14014.0316/P  
; CURRENT APPLICATION NUMBER: US/10/335,394  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/646,028  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 60/077,745  
; PRIOR FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 15  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct  
US-10-335-394-15  
Query Match 86.8%; Score 488; DB 12; Length 374;  
Best Local Similarity 84.9%; Pred. No. 7.1e-37;  
Matches 90; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 DIQMTQSPSSLSASVGRVITCKTSQDINKYMWYQQTGKAPRLLIHYTSLALQGPIS 60  
Db 95 DIQMTQSPSSLSASLGKVIITCKASQDINKYIAWYQHKGPRLLIHYTSLALQGPIS 154

Qy 61 RFSGSGGRDYTTFTISSLPEDIATYYCLOYDNLWTFGGQTKVEIK 106  
Db 155 RFSGSGGRDYTSFISNLEPEDIATYYCLOYDNLWTFGGGKLEIK 200  
RESULT 4  
US-09-229-200A-11  
; Sequence 11, Application US/09229200A  
; Patent No. US20020099179A1  
; GENERAL INFORMATION:  
; APPLICANT: Jolliffe et al.  
; TITLE OF INVENTION: CD4 Specific Recombinant Antibody  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Johnson & Johnson  
; STREET: One Johnson & Johnson Plaza  
; CITY: New Brunswick  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08933-7003  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: <Unknown>  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/229,200A  
; FILING DATE: 13-Jan-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: John W. Wallen, III  
; REGISTRATION NUMBER: 35,403  
; REFERENCE/DOCKET NUMBER: ORT-948  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (858) 784-3239  
; TELEFAX: (908) 524-2808  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-229-200A-11  
Query Match 85.3%; Score 479.5; DB 9; Length 108;  
Best Local Similarity 85.8%; Pred. No. 1.2e-36;  
Matches 91; Conservative 7; Mismatches 7; Indels 1; Gaps 1;  
Qy 1 DIQMTQSPSSLSASVGRVITCKTSQDINKYMWYQQTGKAPRLLIHYTSLALQGPIS 60  
Db 1 DIQMTQSPSSLSASVGRVITCKASPDINNYQQTGKAPRLLIHYTSLALQGPIS 60  
Qy 61 RFSGSGGRDYTTFTISSLPEDIATYYCLOYDNLWTFGGQTKVEI 105  
Db 61 RFSGSGGRDYTTFTISSLPEDIATYYCQQYDNLIFTFGGQTKLIQ 106  
RESULT 5  
US-09-229-200A-15  
; Sequence 15, Application US/09229200A  
; Patent No. US20020099179A1  
; GENERAL INFORMATION:  
; APPLICANT: Jolliffe et al.  
; TITLE OF INVENTION: CD4 Specific Recombinant Antibody  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Johnson & Johnson  
; STREET: One Johnson & Johnson Plaza  
; CITY: New Brunswick  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08933-7003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: November 21, 2003, 15:21:14 ; Search time 19.6027 Seconds  
(without alignments)  
520.023 Million cell updates/sec

Title: US-09-010-377-1

Perfect score: 562

Sequence: 1 DIQMTQSPSSLSASVGRVT.....YCLQYDNLWTFGGTKVEIK 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	86.8	104	2	S26330
2	479	85.2	104	2	S26329
3	478	85.1	106	2	C3936
4	463.5	82.5	125	2	S0365
5	452	80.4	97	2	PH1064
6	446.5	79.4	107	2	PL0270
7	443.5	78.9	108	1	K1HURE
8	443.5	78.9	129	2	S52789
9	443	78.8	103	2	S26332
10	440.5	78.4	107	2	PL0272
11	440	78.3	94	2	E33730
12	439.5	78.2	107	2	PL0269
13	439.5	78.2	107	2	PL0271
14	439.5	78.2	108	1	K1HUAU
15	435.5	77.5	108	2	I39154
16	429.5	76.4	110	2	S44118
17	429.5	76.4	127	2	S40367
18	426	75.8	107	2	S40375
19	424.5	75.5	125	2	S40333
20	421.5	75.0	108	1	K1HULY
21	419.5	74.6	108	1	K1HUAU
22	419.5	74.6	108	1	K1HURU
23	418.5	74.5	123	2	S40331
24	415.5	73.9	109	2	S31998
25	414.5	73.8	131	2	S40352
26	414	73.7	124	2	S40336
27	413.5	73.6	108	1	K1HUSW
28	412.5	73.4	108	2	B49047
29	412	73.3	106	2	PC2397

anti-tetanus toxin

30 410.5 73.0 108 2 S44122 Ig kappa chain V r  
31 410.5 73.0 109 2 S31981 Ig kappa chain - h  
32 409.5 72.9 129 1 K1HUKK Ig kappa chain pre  
33 409 72.8 107 1 K1HUAU Ig kappa chain V-I  
34 408.5 72.7 108 1 K1HURE Ig kappa chain V-I  
35 408.5 72.7 139 2 S40365 Ig kappa chain - h  
36 407.5 72.5 141 2 A49134 Ig kappa chain V-I  
37 405.5 72.2 117 2 S46371 Ig kappa chain V-J  
38 405 72.1 108 2 S30521 Ig kappa chain V r  
39 404.5 72.0 132 2 S40334 Ig kappa chain - h  
40 403.5 71.8 94 2 PH1063 Ig light chain V r  
41 403.5 71.8 107 2 S36264 Ig lambda chain V  
42 403.5 71.8 127 2 S11240 Ig kappa chain V r  
43 403 71.7 117 2 S43528 Ig kappa chain V r  
44 403 71.7 117 2 S42263 Ig kappa chain V r  
45 402.5 71.6 108 1 K1HURU Ig kappa chain V-I

#### ALIGNMENTS

##### RESULT 1

S26330

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000

C:Accession: S26330

R:Stark, S.E.; Caton, A.J.

J. Exp. Med. 174, 613-624, 1991

A:Title: Antibodies that are specific for a single amino acid interchange in a protein ei  
A:Reference number: S26309; MUID:91341421; PMID:1908510

A:Accession: S26330

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-104 <STA>

A:Cross-references: EMBL:X59185; NID:G52316; PIDN:CAA41895.1; PID:gl334063

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 86.8%; Score 488; DB 2; Length 104;

Best Local Similarity 85.6%; Pred. No. 1.6e-35;

Matches 89; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIQMTQSPSSLSASVGRVTITCKTSQDINKYMWYQOTPGKAPRLLIHYTSALQGPIS 60

Db 1 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPELLIHYTSTLQGPIS 60

QY 61 RESGSGGRDYTFITSSLPEDIAITYYCLQYDNLWTFGGTKVE 104

Db 61 RESGSGGRDYTFITSSLPEDIAITYYCLQYDNLWTFGGTKLE 104

##### RESULT 2

S26329

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000

C:Accession: S26329

R:Stark, S.E.; Caton, A.J.

J. Exp. Med. 174, 613-624, 1991

A:Title: Antibodies that are specific for a single amino acid interchange in a protein ei  
A:Reference number: S26309; MUID:91341421; PMID:1908510

A:Accession: S26329

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-104 <STA>

A:Cross-references: EMBL:X59173; NID:G52309; PIDN:CAA41883.1; PID:gl334059

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-90/Domain: immunoglobulin homology <IMM>

Query Match

85.2%; Score 479; DB 2; Length 104;

```

Best Local Similarity 84.6%; Pred. No. 9.8e-35;
Matches 88; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DIQMTQSPSSLSASVGRVITTCSTQDINKYMAWYQOTPGKAPRLIIHYTSLQPGIPS 60
Db 1 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPRLLIIHYTSLQPGIPS 60

Qy 61 RFGSGSGRDYFTTISLQPEDIATYCYQVDNLWTFGGTKVE 104
Db 61 RFGSGSGRDYFTSISNLEPEDIATYCYQVDNLWTFGGTKLE 104

RESULT 3
C33936
Ig kappa chain V region (VIM113) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 21-Jan-2000
C:Accession: C33936
A:Title: Nucleotide changes in sequential variants of influenza virus hemagglutinin gene
A:Reference number: A33936; MUID:89282831; PMID:2471975
A:Accession: C33936
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-106 <MEE>
A:Cross-references: GB:J04577; NID:9623187; PIDN:AAA60443.1; PID:9623189
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 85.1%; Score 478; DB 2; Length 106;
Best Local Similarity 83.0%; Pred. No. 1.2e-34;
Matches 88; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DIQMTQSPSSLSASVGRVITTCSTQDINKYMAWYQOTPGKAPRLIIHYTSLQPGIPS 60
Db 1 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPRLLIIHYTSLQPGIPS 60

Qy 61 RFGSGSGRDYFTTISLQPEDIATYCYQVDNLWTFGGTKVEIK 106
Db 61 RFGSGSGRDYFTSISNLDPEEATYCYQVDNLWTFGGTKLEIK 106

RESULT 4
S09365
Ig kappa chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S09365
R:Pedersen, R.; van Ness, B.
Nucleic Acids Res. 17, 9797-9809, 1989
A:Title: Direct evidence for intrastand DNA inversion of kappa immunoglobulin gene segm
A:Reference number: S09365; MUID:90098844; PMID:2513557
A:Accession: S09365
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <FED>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:33-107/Domain: immunoglobulin homology <IMM>

Query Match 82.5%; Score 463.5; DB 2; Length 125;
Best Local Similarity 83.2%; Pred. No. 2.6e-33;
Matches 89; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

Qy 1 DIQMTQSPSSLSASVGRVITTCSTQDINKYMAWYQOTPGKAPRLIIHYTSLQPGIPS 60
Db 18 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPRLLIIHYTSLQPGIPS 77

Qy 61 RFGSGSGRDYFTTISLQPEDIATYCYQVDNLWTFGGTKVEIK 106
Db 78 RFGSGSGSDYFTSISNLDPEEATYCYQVDNLWTFGGTKLEIK 124

```

```

RESULT 5
PH1064
Ig light chain V region (clone 202.54) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PH1064
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B ce
A:Reference number: PH0971; MUID:92381444; PMID:1512540
A:Accession: PH1064
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-97 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 80.4%; Score 452; DB 2; Length 97;
Best Local Similarity 85.6%; Pred. No. 2e-32;
Matches 83; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DIQMTQSPSSLSASVGRVITTCSTQDINKYMAWYQOTPGKAPRLIIHYTSLQPGIPS 60
Db 1 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPRLLIIHYTSLQPGIPS 60

Qy 61 RFGSGSGRDYFTTISLQPEDIATYCYQVDNLWTF 97
Db 61 RFGSGSGRDYFTSISNLEPEDIATYCYQVDNLWTF 97

RESULT 6
PL0270
Ig kappa chain V region (anti-DNA, 6G6VK) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C:Accession: PL0270
R:Shlomchik, M.; Mascellini, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.
J. Exp. Med. 171, 265-297, 1990
A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic m
A:Reference number: PL0231; MUID:90111618; PMID:2104919
A:Accession: PL0270
A:Molecule type: mRNA
A:Residues: 1-107 <SHL>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>
F:24-34/Region: complementarity-determining 1
F:35-49/Region: framework 2
F:50-56/Region: complementarity-determining 2
F:57-88/Region: framework 3
F:89-97/Region: complementarity-determining 3
F:98-107/Region: framework 4

Query Match 79.4%; Score 446.5; DB 2; Length 107;
Best Local Similarity 79.4%; Pred. No. 6.5e-32;
Matches 85; Conservative 11; Mismatches 10; Indels 1; Gaps 1;

Qy 1 DIQMTQSPSSLSASVGRVITTCSTQDINKYMAWYQOTPGKAPRLIIHYTSLQPGIPS 60
Db 1 DIQMTQSPSSLSASLGKVTITCKASQDINKYIAWYQHKPGKPRLLIIHYTSLQPGIPS 60

Qy 61 RFGSGSGRDYFTTISLQPEDIATYCYQVDNLWTFGGTKVEIK 106
Db 61 RFGSGSGRDYFTSISNLEPEDIATYCYQVDNLWTFGGTKLEIK 107

RESULT 7
KHURE

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

un on: November 21, 2003, 15:20:03 ; Search time 10.5274 Seconds  
(without alignments)  
473.510 Million cell updates/sec

File: US-09-010-377-1

Perfect score: 562

Sequence: 1 DIQWTQSPSLASVGRVT.....YCLQYDNLWTFGQTKVEIK 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	443.5	78.9	108	1 KV10 HUMAN	P01607 homo sapien
2	439.5	78.2	108	1 KV1B HUMAN	P01594 homo sapien
3	421.5	75.0	108	1 KV1A HUMAN	P01605 homo sapien
4	419.5	74.6	108	1 KV1A HUMAN	P01593 homo sapien
5	419.5	74.6	108	1 KV1P HUMAN	P01608 homo sapien
6	419.5	74.6	108	1 KV1V HUMAN	P80362 homo sapien
7	413.5	73.6	108	1 KV1Q HUMAN	P01609 homo sapien
8	409.5	72.9	129	1 KV1Q HUMAN	P04431 homo sapien
9	409	72.8	107	1 KV1D HUMAN	P01596 homo sapien
10	408.5	72.7	108	1 KV1H HUMAN	P01610 homo sapien
11	402.5	71.6	108	1 KV1H HUMAN	P01600 homo sapien
12	397.5	70.7	108	1 KV1F HUMAN	P01598 homo sapien
13	396.5	70.6	108	1 KV1F HUMAN	P04430 homo sapien
14	390.5	69.5	108	1 KV1K HUMAN	P01603 homo sapien
15	388.5	69.1	108	1 KV1E HUMAN	P01597 homo sapien
16	386.5	68.8	108	1 KV1L HUMAN	P01604 homo sapien
17	385.5	68.6	108	1 KV1N HUMAN	P01606 homo sapien
18	384.5	68.4	108	1 KV1J HUMAN	P01595 homo sapien
19	383.5	68.2	108	1 KV1G HUMAN	P01599 homo sapien
20	383.5	68.2	129	1 KV1S HUMAN	P04432 homo sapien
21	381.5	67.9	108	1 KV1S HUMAN	P01611 homo sapien
22	376.5	67.0	134	1 KV4C HUMAN	P06314 homo sapien
23	371.5	66.1	108	1 KV5T MOUSE	P01643 mus musculu
24	368	65.5	117	1 KV1J HUMAN	P01602 homo sapien
25	366.5	65.2	128	1 KV5E MOUSE	P01637 mus musculu
26	365	64.9	133	1 KV4B HUMAN	P06313 homo sapien
27	364.5	64.9	108	1 KV5O MOUSE	P01648 mus musculu
28	363.5	64.7	108	1 KV5N MOUSE	P01647 mus musculu
29	362	64.4	109	1 KV1T HUMAN	P01612 homo sapien
30	362	64.4	117	1 KV1J HUMAN	P01601 homo sapien
31	361.5	64.3	108	1 KV5K MOUSE	P01644 mus musculu
32	359.5	64.0	108	1 KV5M MOUSE	P01646 mus musculu
33	358.5	63.8	108	1 KV5L MOUSE	P01645 mus musculu

ALIGNMENTS

RESULT 1

ID	KV10_HUMAN	STANDARD;	PRT;	108 AA.
AC	P01607;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig kappa chain V-1 region Rei.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI TaxID=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=76023758; PubMed=809329;			
RA	Palm W., Hilschmann N.;			
RT	"The primary structure of a crystalline monoclonal immunoglobulin			
RT	kappa-type L-chain, subgroup I (Bence-Jones protein Rei); isolation			
RT	and characterization of the tryptic peptides; the complete amino acid			
RT	sequence of the protein; a contribution to the elucidation of the			
RT	three-dimensional structure of antibodies, in particular their			
RT	combining site.";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 356:167-191(1975).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).			
RX	MEDLINE=76039968; PubMed=1182131;			
RA	Epp O., Lattman E.E., Schiffer N., Huber R., Palm W.;			
RT	"The molecular structure of a dimer composed of the variable portions			
RT	of the Bence-Jones protein REI refined at 2.0-A resolution.";			
RL	Biochemistry 14:4943-4952(1975).			
CC	!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)			
CC	MARKER.			
CC	!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.			
DR	PIR; A91663; KIHURE.			
DR	PDB; 1REI; 17-FEB-84.			
DR	PDB; 1AR2; 12-NOV-97.			
DR	PDB; 1BWV; 29-DEC-99.			
DR	GO; GO:0005576; C:extracellular; NAS.			
DR	GO; GO:0003823; F:antigen binding activity; NAS.			
DR	GO; GO:0006955; P:immune response; NAS.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig MHC.			
DR	InterPro; IPR003596; Ig_v.			
DR	Pfam; PF00047; ig; 1.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
DR	Immunoglobulin V region; Bence-Jones protein; 3D-structure.			
KW	DOMAIN 1 23			
FT	DOMAIN 24 34			
FT	COMPLEMENTARITY-DETERMINING-1.			
FT	DOMAIN 35 49			
FT	COMPLEMENTARITY-DETERMINING-2.			
FT	DOMAIN 50 56			
FT	COMPLEMENTARITY-DETERMINING-3.			
FT	DOMAIN 57 88			
FT	COMPLEMENTARITY-DETERMINING-3.			
FT	DOMAIN 89 97			
FT	COMPLEMENTARITY-DETERMINING-3.			
FT	DOMAIN 98 107			
FT	DISULFID 23 88			
FT	STRAND 4 7			

P01613 homo sapien  
P01625 homo sapien  
P04207 homo sapien  
P18136 homo sapien  
P01653 mus musculu  
P01638 mus musculu  
P01650 mus musculu  
P01623 homo sapien  
P01652 mus musculu  
P04946 mus sapien  
P01622 homo sapien  
P01624 homo sapien

```

STRAND      10 13
TURN        15 16
STRAND      19 25
TURN        30 31
STRAND      33 38
TURN        40 41
STRAND      44 49
TURN        50 52
STRAND      53 54
TURN        56 57
STRAND      60 61
TURN        62 67
STRAND      68 69
TURN        70 75
HELIX       80 82
STRAND      85 90
TURN        97 98
STRAND     102 106
TURN        108 108
NON TER
SEQUENCE 108 AA; 11902 MW; 9E8143E1188BCE2A CRC64;

Query Match      78.9%; Score 443.5; DB 1; Length 108;
Best Local Similarity 81.1%; Pred. No. 4.9e-40;
Matches 86; Conservative 9; Mismatches 10; Indels 1; Gaps 1;

QY 1 DIQMTQSPSSLSASVGRVTITCKTSQDINKYMWYQQTGKAPRLIIHYTSALQFGIPS 60
DB 1 DIQMTQSPSSLSASVGRVTITCKTSQDINKYMWYQQTGKAPRLIIHYTSALQAGVPS 60

QY 61 RPSGSGSRDYTTFTTSSLPQEDYATYCYQDNL-WTFQGGTKVEI 105
DB 61 RPSGSGSRDYTTFTTSSLPQEDYATYCYQYQSLPYTFQGGTKLQI 106

RESULT 2
KV1B HUMAN STANDARD; PRT; 108 AA.
ID KV1B HUMAN
AC P01594;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig kappa chain V-I region AU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=72189444; PubMed=5028201;
RX Schiechl H., Hilschmann N.;
RA "Rule of antibody structure. The primary structure of a monoclonal
RT immunoglobulin L-chain of the kappa-type, subgroup I (Bence-Jones
RL protein Au).";
RL Hoppe-Seyler's Z. Physiol. Chem. 353:345-370 (1972).
RN [2]
RX X-RAY CRYSTALLOGRAPHY.
RP MEDLINE=77022433; PubMed=1234024;
RX Fehllhammer H., Schiffer M., Epp O., Colman P.M., Lattman E.E.,
RA Schwager P., Steigemann W., Schramm H.J.;
RT "The structure determination of the variable portion of the
RL Bence-Jones protein Au.";
RN Biophys. Struct. Mech. 1:139-146 (1975).
CC -I- MISCELLANEOUS: THE STRUCTURE OF THE V REGION WAS DETERMINED BY
CC MOLECULAR REPLACEMENT METHODS USING THE KNOWN STRUCTURE OF THE V
CC REGION OF THE KAPPA CHAIN REI.
CC -I- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -I- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PDB; 1JVS; 30-JAN-02.
DR A91653; KIHUAU.
DR GO:0005576; C:extracellular; NAS.
DR GO:0003823; F:antigen binding activity; NAS.
DR GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.

```

```

DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region; Bence-Jones protein; 3D-structure.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON TER 108 108
SQ SEQUENCE 108 AA; 11939 MW; E8011187EE6F6FB9 CRC64;

Query Match      78.2%; Score 439.5; DB 1; Length 108;
Best Local Similarity 79.4%; Pred. No. 1.3e-39;
Matches 85; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY 1 DIQMTQSPSSLSASVGRVTITCKTSQDINKYMWYQQTGKAPRLIIHYTSALQFGIPS 60
DB 1 DIQMTQSPSSLSASVGRVTITCKTSQDINKYMWYQQTGKAPRLIIHYTSALQAGVPS 60

QY 61 RPSGSGSRDYTTFTTSSLPQEDYATYCYQDNL-WTFQGGTKVEIK 106
DB 61 RPSGSGSRDYTTFTTSSLPQEDYATYCYQYDYLEWTFQGGTKVEIK 107

RESULT 3
KV1M HUMAN STANDARD; PRT; 108 AA.
ID KV1M HUMAN
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=77038198; PubMed=824717;
RX Capra J.D., Klapper D.G.;
RA "Complete amino acid sequence of the variable domains of two human
RT IgM anti-gamma globulins (Lay/Pom) with shared idiotypic
RT specificities.";
RL Scand. J. Immunol. 5:677-684 (1976).
CC -I- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01871; KIHULY.
DR HSSP; P01607; IREI.
DR GO:0005576; C:extracellular; NAS.
DR GO:0003823; F:antigen binding activity; NAS.
DR GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 99
FT DOMAIN 100 107
FT DOMAIN 108 108
FT DOMAIN 109 109
FT DOMAIN 110 110
FT DOMAIN 111 110
FT DOMAIN 112 110
FT DOMAIN 113 110
FT DOMAIN 114 110
FT DOMAIN 115 110
FT DOMAIN 116 110
FT DOMAIN 117 110
FT DOMAIN 118 110
FT DOMAIN 119 110
FT DOMAIN 120 110
FT DOMAIN 121 110
FT DOMAIN 122 110
FT DOMAIN 123 110
FT DOMAIN 124 110
FT DOMAIN 125 110
FT DOMAIN 126 110
FT DOMAIN 127 110
FT DOMAIN 128 110
FT DOMAIN 129 110
FT DOMAIN 130 110
FT DOMAIN 131 110
FT DOMAIN 132 110
FT DOMAIN 133 110
FT DOMAIN 134 110
FT DOMAIN 135 110
FT DOMAIN 136 110
FT DOMAIN 137 110
FT DOMAIN 138 110
FT DOMAIN 139 110
FT DOMAIN 140 110
FT DOMAIN 141 110
FT DOMAIN 142 110
FT DOMAIN 143 110
FT DOMAIN 144 110
FT DOMAIN 145 110
FT DOMAIN 146 110
FT DOMAIN 147 110
FT DOMAIN 148 110
FT DOMAIN 149 110
FT DOMAIN 150 110
FT DOMAIN 151 110
FT DOMAIN 152 110
FT DOMAIN 153 110
FT DOMAIN 154 110
FT DOMAIN 155 110
FT DOMAIN 156 110
FT DOMAIN 157 110
FT DOMAIN 158 110
FT DOMAIN 159 110
FT DOMAIN 160 110
FT DOMAIN 161 110
FT DOMAIN 162 110
FT DOMAIN 163 110
FT DOMAIN 164 110
FT DOMAIN 165 110
FT DOMAIN 166 110
FT DOMAIN 167 110
FT DOMAIN 168 110
FT DOMAIN 169 110
FT DOMAIN 170 110
FT DOMAIN 171 110
FT DOMAIN 172 110
FT DOMAIN 173 110
FT DOMAIN 174 110
FT DOMAIN 175 110
FT DOMAIN 176 110
FT DOMAIN 177 110
FT DOMAIN 178 110
FT DOMAIN 179 110
FT DOMAIN 180 110
FT DOMAIN 181 110
FT DOMAIN 182 110
FT DOMAIN 183 110
FT DOMAIN 184 110
FT DOMAIN 185 110
FT DOMAIN 186 110
FT DOMAIN 187 110
FT DOMAIN 188 110
FT DOMAIN 189 110
FT DOMAIN 190 110
FT DOMAIN 191 110
FT DOMAIN 192 110
FT DOMAIN 193 110
FT DOMAIN 194 110
FT DOMAIN 195 110
FT DOMAIN 196 110
FT DOMAIN 197 110
FT DOMAIN 198 110
FT DOMAIN 199 110
FT DOMAIN 200 110
FT DOMAIN 201 110
FT DOMAIN 202 110
FT DOMAIN 203 110
FT DOMAIN 204 110
FT DOMAIN 205 110
FT DOMAIN 206 110
FT DOMAIN 207 110
FT DOMAIN 208 110
FT DOMAIN 209 110
FT DOMAIN 210 110
FT DOMAIN 211 110
FT DOMAIN 212 110
FT DOMAIN 213 110
FT DOMAIN 214 110
FT DOMAIN 215 110
FT DOMAIN 216 110
FT DOMAIN 217 110
FT DOMAIN 218 110
FT DOMAIN 219 110
FT DOMAIN 220 110
FT DOMAIN 221 110
FT DOMAIN 222 110
FT DOMAIN 223 110
FT DOMAIN 224 110
FT DOMAIN 225 110
FT DOMAIN 226 110
FT DOMAIN 227 110
FT DOMAIN 228 110
FT DOMAIN 229 110
FT DOMAIN 230 110
FT DOMAIN 231 110
FT DOMAIN 232 110
FT DOMAIN 233 110
FT DOMAIN 234 110
FT DOMAIN 235 110
FT DOMAIN 236 110
FT DOMAIN 237 110
FT DOMAIN 238 110
FT DOMAIN 239 110
FT DOMAIN 240 110
FT DOMAIN 241 110
FT DOMAIN 242 110
FT DOMAIN 243 110
FT DOMAIN 244 110
FT DOMAIN 245 110
FT DOMAIN 246 110
FT DOMAIN 247 110
FT DOMAIN 248 110
FT DOMAIN 249 110
FT DOMAIN 250 110
FT DOMAIN 251 110
FT DOMAIN 252 110
FT DOMAIN 253 110
FT DOMAIN 254 110
FT DOMAIN 255 110
FT DOMAIN 256 110
FT DOMAIN 257 110
FT DOMAIN 258 110
FT DOMAIN 259 110
FT DOMAIN 260 110
FT DOMAIN 261 110
FT DOMAIN 262 110
FT DOMAIN 263 110
FT DOMAIN 264 110
FT DOMAIN 265 110
FT DOMAIN 266 110
FT DOMAIN 267 110
FT DOMAIN 268 110
FT DOMAIN 269 110
FT DOMAIN 270 110
FT DOMAIN 271 110
FT DOMAIN 272 110
FT DOMAIN 273 110
FT DOMAIN 274 110
FT DOMAIN 275 110
FT DOMAIN 276 110
FT DOMAIN 277 110
FT DOMAIN 278 110
FT DOMAIN 279 110
FT DOMAIN 280 110
FT DOMAIN 281 110
FT DOMAIN 282 110
FT DOMAIN 283 110
FT DOMAIN 284 110
FT DOMAIN 285 110
FT DOMAIN 286 110
FT DOMAIN 287 110
FT DOMAIN 288 110
FT DOMAIN 289 110
FT DOMAIN 290 110
FT DOMAIN 291 110
FT DOMAIN 292 110
FT DOMAIN 293 110
FT DOMAIN 294 110
FT DOMAIN 295 110
FT DOMAIN 296 110
FT DOMAIN 297 110
FT DOMAIN 298 110
FT DOMAIN 299 110
FT DOMAIN 300 110
FT DOMAIN 301 110
FT DOMAIN 302 110
FT DOMAIN 303 110
FT DOMAIN 304 110
FT DOMAIN 305 110
FT DOMAIN 306 110
FT DOMAIN 307 110
FT DOMAIN 308 110
FT DOMAIN 309 110
FT DOMAIN 310 110
FT DOMAIN 311 110
FT DOMAIN 312 110
FT DOMAIN 313 110
FT DOMAIN 314 110
FT DOMAIN 315 110
FT DOMAIN 316 110
FT DOMAIN 317 110
FT DOMAIN 318 110
FT DOMAIN 319 110
FT DOMAIN 320 110
FT DOMAIN 321 110
FT DOMAIN 322 110
FT DOMAIN 323 110
FT DOMAIN 324 110
FT DOMAIN 325 110
FT DOMAIN 326 110
FT DOMAIN 327 110
FT DOMAIN 328 110
FT DOMAIN 329 110
FT DOMAIN 330 110
FT DOMAIN 331 110
FT DOMAIN 332 110
FT DOMAIN 333 110
FT DOMAIN 334 110
FT DOMAIN 335 110
FT DOMAIN 336 110
FT DOMAIN 337 110
FT DOMAIN 338 110
FT DOMAIN 339 110
FT DOMAIN 340 110
FT DOMAIN 341 110
FT DOMAIN 342 110
FT DOMAIN 343 110
FT DOMAIN 344 110
FT DOMAIN 345 110
FT DOMAIN 346 110
FT DOMAIN 347 110
FT DOMAIN 348 110
FT DOMAIN 349 110
FT DOMAIN 350 110
FT DOMAIN 351 110
FT DOMAIN 352 110
FT DOMAIN 353 110
FT DOMAIN 354 110
FT DOMAIN 355 110
FT DOMAIN 356 110
FT DOMAIN 357 110
FT DOMAIN 358 110
FT DOMAIN 359 110
FT DOMAIN 360 110
FT DOMAIN 361 110
FT DOMAIN 362 110
FT DOMAIN 363 110
FT DOMAIN 364 110
FT DOMAIN 365 110
FT DOMAIN 366 110
FT DOMAIN 367 110
FT DOMAIN 368 110
FT DOMAIN 369 110
FT DOMAIN 370 110
FT DOMAIN 371 110
FT DOMAIN 372 110
FT DOMAIN 373 110
FT DOMAIN 374 110
FT DOMAIN 375 110
FT DOMAIN 376 110
FT DOMAIN 377 110
FT DOMAIN 378 110
FT DOMAIN 379 110
FT DOMAIN 380 110
FT DOMAIN 381 110
FT DOMAIN 382 110
FT DOMAIN 383 110
FT DOMAIN 384 110
FT DOMAIN 385 110
FT DOMAIN 386 110
FT DOMAIN 387 110
FT DOMAIN 388 110
FT DOMAIN 389 110
FT DOMAIN 390 110
FT DOMAIN 391 110
FT DOMAIN 392 110
FT DOMAIN 393 110
FT DOMAIN 394 110
FT DOMAIN 395 110
FT DOMAIN 396 110
FT DOMAIN 397 110
FT DOMAIN 398 110
FT DOMAIN 399 110
FT DOMAIN 400 110
FT DOMAIN 401 110
FT DOMAIN 402 110
FT DOMAIN 403 110
FT DOMAIN 404 110
FT DOMAIN 405 110
FT DOMAIN 406 110
FT DOMAIN 407 110
FT DOMAIN 408 110
FT DOMAIN 409 110
FT DOMAIN 410 110
FT DOMAIN 411 110
FT DOMAIN 412 110
FT DOMAIN 413 110
FT DOMAIN 414 110
FT DOMAIN 415 110
FT DOMAIN 416 110
FT DOMAIN 417 110
FT DOMAIN 418 110
FT DOMAIN 419 110
FT DOMAIN 420 110
FT DOMAIN 421 110
FT DOMAIN 422 110
FT DOMAIN 423 110
FT DOMAIN 424 110
FT DOMAIN 425 110
FT DOMAIN 426 110
FT DOMAIN 427 110
FT DOMAIN 428 110
FT DOMAIN 429 110
FT DOMAIN 430 110
FT DOMAIN 431 110
FT DOMAIN 432 110
FT DOMAIN 433 110
FT DOMAIN 434 110
FT DOMAIN 435 110
FT DOMAIN 436 110
FT DOMAIN 437 110
FT DOMAIN 438 110
FT DOMAIN 439 110
FT DOMAIN 440 110
FT DOMAIN 441 110
FT DOMAIN 442 110
FT DOMAIN 443 110
FT DOMAIN 444 110
FT DOMAIN 445 110
FT DOMAIN 446 110
FT DOMAIN 447 110
FT DOMAIN 448 110
FT DOMAIN 449 110
FT DOMAIN 450 110
FT DOMAIN 451 110
FT DOMAIN 452 110
FT DOMAIN 453 110
FT DOMAIN 454 110
FT DOMAIN 455 110
FT DOMAIN 456 110
FT DOMAIN 457 110
FT DOMAIN 458 110
FT DOMAIN 459 110
FT DOMAIN 460 110
FT DOMAIN 461 110
FT DOMAIN 462 110
FT DOMAIN 463 110
FT DOMAIN 464 110
FT DOMAIN 465 110
FT DOMAIN 466 110
FT DOMAIN 467 110
FT DOMAIN 468 110
FT DOMAIN 469 110
FT DOMAIN 470 110
FT DOMAIN 471 110
FT DOMAIN 472 110
FT DOMAIN 473 110
FT DOMAIN 474 110
FT DOMAIN 475 110
FT DOMAIN 476 110
FT DOMAIN 477 110
FT DOMAIN 478 110
FT DOMAIN 479 110
FT DOMAIN 480 110
FT DOMAIN 481 110
FT DOMAIN 482 110
FT DOMAIN 483 110
FT DOMAIN 484 110
FT DOMAIN 485 110
FT DOMAIN 486 110
FT DOMAIN 487 110
FT DOMAIN 488 110
FT DOMAIN 489 110
FT DOMAIN 490 110
FT DOMAIN 491 110
FT DOMAIN 492 110
FT DOMAIN 493 110
FT DOMAIN 494 110
FT DOMAIN 495 110
FT DOMAIN 496 110
FT DOMAIN 497 110
FT DOMAIN 498 110
FT DOMAIN 499 110
FT DOMAIN 500 110
FT DOMAIN 501 110
FT DOMAIN 502 110
FT DOMAIN 503 110
FT DOMAIN 504 110
FT DOMAIN 505 110
FT DOMAIN 506 110
FT DOMAIN 507 110
FT DOMAIN 508 110
FT DOMAIN 509 110
FT DOMAIN 510 110
FT DOMAIN 511 110
FT DOMAIN 512 110
FT DOMAIN 513 110
FT DOMAIN 514 110
FT DOMAIN 515 110
FT DOMAIN 516 110
FT DOMAIN 517 110
FT DOMAIN 518 110
FT DOMAIN 519 110
FT DOMAIN 520 110
FT DOMAIN 521 110
FT DOMAIN 522 110
FT DOMAIN 523 110
FT DOMAIN 524 110
FT DOMAIN 525 110
FT DOMAIN 526 110
FT DOMAIN 527 110
FT DOMAIN 528 110
FT DOMAIN 529 110
FT DOMAIN 530 110
FT DOMAIN 531 110
FT DOMAIN 532 110
FT DOMAIN 533 110
FT DOMAIN 534 110
FT DOMAIN 535 110
FT DOMAIN 536 110
FT DOMAIN 537 110
FT DOMAIN 538 110
FT DOMAIN 539 110
FT DOMAIN 540 110
FT DOMAIN 541 110
FT DOMAIN 542 110
FT DOMAIN 543 110
FT DOMAIN 544 110
FT DOMAIN 545 110
FT DOMAIN 546 110
FT DOMAIN 547 110
FT DOMAIN 548 110
FT DOMAIN 549 110
FT DOMAIN 550 110
FT DOMAIN 551 110
FT DOMAIN 552 110
FT DOMAIN 553 110
FT DOMAIN 554 110
FT DOMAIN 555 110
FT DOMAIN 556 110
FT DOMAIN 557 110
FT DOMAIN 558 110
FT DOMAIN 559 110
FT DOMAIN 560 110
FT DOMAIN 561 110
FT DOMAIN 562 110
FT DOMAIN 563 110
FT DOMAIN 564 110
FT DOMAIN 565 110
FT DOMAIN 566 110
FT DOMAIN 567 110
FT DOMAIN 568 110
FT DOMAIN 569 110
FT DOMAIN 570 110
FT DOMAIN 571 110
FT DOMAIN 572 110
FT DOMAIN 573 110
FT DOMAIN 574 110
FT DOMAIN 575 110
FT DOMAIN 576 110
FT DOMAIN 577 110
FT DOMAIN 578 110
FT DOMAIN 579 110
FT DOMAIN 580 110
FT DOMAIN 581 110
FT DOMAIN 582 110
FT DOMAIN 583 110
FT DOMAIN 584 110
FT DOMAIN 585 110
FT DOMAIN 586 110
FT DOMAIN 587 110
FT DOMAIN 588 110
FT DOMAIN 589 110
FT DOMAIN 590 110
FT DOMAIN 591 110
FT DOMAIN 592 110
FT DOMAIN 593 110
FT DOMAIN 594 110
FT DOMAIN 595 110
FT DOMAIN 596 110
FT DOMAIN 597 110
FT DOMAIN 598 110
FT DOMAIN 599 110
FT DOMAIN 600 110
FT DOMAIN 601 110
FT DOMAIN 602 110
FT DOMAIN 603 110
FT DOMAIN 604 110
FT DOMAIN 605 110
FT DOMAIN 606 110
FT DOMAIN 607 110
FT DOMAIN 608 110
FT DOMAIN 609 110
FT DOMAIN 610 110
FT DOMAIN 611 110
FT DOMAIN 612 110
FT DOMAIN 613 110
FT DOMAIN 614 110
FT DOMAIN 615 110
FT DOMAIN 616 110
FT DOMAIN 617 110
FT DOMAIN 618 110
FT DOMAIN 619 110
FT DOMAIN 620 110
FT DOMAIN 621 110
FT DOMAIN 622 110
FT DOMAIN 623 110
FT DOMAIN 624 110
FT DOMAIN 625 110
FT DOMAIN 626 110
FT DOMAIN 627 110
FT DOMAIN 628 110
FT DOMAIN 629 110
FT DOMAIN 630 110
FT DOMAIN 631 110
FT DOMAIN 632 110
FT DOMAIN 633 110
FT DOMAIN 634 110
FT DOMAIN 635 110
FT DOMAIN 636 110
FT DOMAIN 637 110
FT DOMAIN 638 110
FT DOMAIN 639 110
FT DOMAIN 640 110
FT DOMAIN 641 110
FT DOMAIN 642 110
FT DOMAIN 643 110
FT DOMAIN 644 110
FT DOMAIN 645 110
FT DOMAIN 646 110
FT DOMAIN 647 110
FT DOMAIN 648 110
FT DOMAIN 649 110
FT DOMAIN 650 110
FT DOMAIN 651 110
FT DOMAIN 652 110
FT DOMAIN 653 110
FT DOMAIN 654 110
FT DOMAIN 655 110
FT DOMAIN 656 110
FT DOMAIN 657 110
FT DOMAIN 658 110
FT DOMAIN 659 110
FT DOMAIN 660 110
FT DOMAIN 661 110
FT DOMAIN 662 110
FT DOMAIN 663 110
FT DOMAIN 664 110
FT DOMAIN 665 110
FT DOMAIN 666 110
FT DOMAIN 667 110
FT DOMAIN 668 110
FT DOMAIN 669 110
FT DOMAIN 670 110
FT DOMAIN 671 110
FT DOMAIN 672 110
FT DOMAIN 673 110
FT DOMAIN 674 110
FT DOMAIN 675 110
FT DOMAIN 676 110
FT DOMAIN 677 110
FT DOMAIN 678 110
FT DOMAIN 679 110
FT DOMAIN 680 110
FT DOMAIN 681 110
FT DOMAIN 682 110
FT DOMAIN 683 110
FT DOMAIN 684 110
FT DOMAIN 685 110
FT DOMAIN 686 110
FT DOMAIN 687 110
FT DOMAIN 688 110
FT DOMAIN 689 110
FT DOMAIN 690 110
FT DOMAIN 691 110
FT DOMAIN 692 110
FT DOMAIN 693 110
FT DOMAIN 694 110
FT DOMAIN 695 110
FT DOMAIN 696 110
FT DOMAIN 697 110
FT DOMAIN 698 110
FT DOMAIN 699 110
FT DOMAIN 700 110
FT DOMAIN 701 110
FT DOMAIN 702 110
FT DOMAIN 703 110
FT DOMAIN 704 110
FT DOMAIN 705 110
FT DOMAIN 706 110
FT DOMAIN 707 110
FT DOMAIN 708 110
FT DOMAIN 709 110
FT DOMAIN 710 110
FT DOMAIN 711 110
FT DOMAIN 712 110
FT DOMAIN 713 110
FT DOMAIN 714 110
FT DOMAIN 715 110
FT DOMAIN 716 110
FT DOMAIN 717 110
FT DOMAIN 718 110
FT DOMAIN 719 110
FT DOMAIN 720 110
FT DOMAIN 721 110
FT DOMAIN 722 110
FT DOMAIN 723 110
FT DOMAIN 724 110
FT DOMAIN 725 110
FT DOMAIN 726 110
FT DOMAIN 727 110
FT DOMAIN 728 110
FT DOMAIN 729 110
FT DOMAIN 730 110
FT DOMAIN 731 110
FT DOMAIN 732 110
FT DOMAIN 733 110
FT DOMAIN 734 110
FT DOMAIN 735 110
FT DOMAIN 736 110
FT DOMAIN 737 110
FT DOMAIN 738 110
FT DOMAIN 739 110
FT DOMAIN 740 110
FT DOMAIN 741 110
FT DOMAIN 742 110
FT DOMAIN 743 110
FT DOMAIN 744 110
FT DOMAIN 745 110
FT DOMAIN 746 110
FT DOMAIN 747 110
FT DOMAIN 748 110
FT DOMAIN 749 110
FT DOMAIN 750 110
FT DOMAIN 751 110
FT DOMAIN 752 110
FT DOMAIN 753 110
FT DOMAIN 754 110
FT DOMAIN 755 110
FT DOMAIN 756 110
FT DOMAIN 757 110
FT DOMAIN 758 110
FT DOMAIN 759 110
FT DOMAIN 760 110
FT DOMAIN 761 110
FT DOMAIN 762 110
FT DOMAIN 763 110
FT DOMAIN 764 110
FT DOMAIN 765 110
FT DOMAIN 766 110
FT DOMAIN 767 110
FT DOMAIN 768 110
FT DOMAIN 769 110
FT DOMAIN 770 110
FT DOMAIN 771 110
FT DOMAIN 772 110
FT DOMAIN 773 110
FT DOMAIN 774 110
FT DOMAIN 775 110
FT DOMAIN 776 110
FT DOMAIN 777 110
FT DOMAIN 778 110
FT DOMAIN 779 110
FT DOMAIN 780 110
FT DOMAIN 781 110
FT DOMAIN 782 110
FT DOMAIN 783 110
FT DOMAIN 784 110
FT DOMAIN 785 110
FT DOMAIN 786 110
FT DOMAIN 787 110
FT DOMAIN 788 110
FT DOMAIN 789 110
FT DOMAIN 790 110
FT DOMAIN 791 110
FT DOMAIN 792 110
FT DOMAIN 793 110
FT DOMAIN 794 110
FT DOMAIN 795 110
FT DOMAIN 796 110
FT DOMAIN 797 110
FT DOMAIN 798 110
FT DOMAIN 799 110
FT DOMAIN 800 110
FT DOMAIN 801 110
FT DOMAIN 802 110
FT DOMAIN 803 110
FT DOMAIN 804 110
FT DOMAIN 805 110
FT DOMAIN 806 110
FT DOMAIN 807 110
FT DOMAIN 808 110
FT DOMAIN 809 110
FT DOMAIN 810 110
FT DOMAIN 811 110
FT DOMAIN 812 110
FT DOMAIN 813 110
FT DOMAIN 814 110
FT DOMAIN 815 110
FT DOMAIN 816 110
FT DOMAIN 817 110
FT DOMAIN 818 110
FT DOMAIN 819 110
FT DOMAIN 820 110
FT DOMAIN 821 110
FT DOMAIN 822 110
FT DOMAIN 823 110
FT DOMAIN 824 110
FT DOMAIN 825 110
FT DOMAIN 826 110
FT DOMAIN 827 110
FT DOMAIN 828 110
FT DOMAIN 829 110
FT DOMAIN 830 110
FT DOMAIN 831 110
FT DOMAIN 832 110
FT DOMAIN 833 110
FT DOMAIN 834 110
FT DOMAIN 835 110
FT DOMAIN 836 110
FT DOMAIN 837 110
FT DOMAIN 838 110
FT DOMAIN 839 110
FT DOMAIN 840 110
FT DOMAIN 841 110
FT DOMAIN 842 110
FT DOMAIN 843 110
FT DOMAIN 844 110
FT DOMAIN 845 110
FT DOMAIN 846 110
FT DOMAIN 847 110
FT DOMAIN 848 110
FT DOMAIN 849 110
FT DOMAIN 850 110
FT DOMAIN 851 110
FT DOMAIN 852 110
FT DOMAIN 853 110
FT DOMAIN 854 110
FT DOMAIN 855 110
FT DOMAIN 856 110
FT DOMAIN 857 110
FT DOMAIN 858 110
FT DOMAIN 859 110
FT DOMAIN 860 110
FT DOMAIN 861 110
FT DOMAIN 862 110
FT DOMAIN 863 110
FT DOMAIN 864 110
FT DOMAIN 865 110
FT DOMAIN 866 110
FT DOMAIN 867 110
FT DOMAIN 868 110
FT DOMAIN 869 110
FT DOMAIN 870 110
FT DOMAIN 871 110
FT DOMAIN 872 110
FT DOMAIN 873 110
FT DOMAIN 874 110
FT DOMAIN 875 110
FT DOMAIN 876 110
FT DOMAIN 877 110
FT DOMAIN 878 110
FT DOMAIN 879 110
FT DOMAIN 880 110
FT DOMAIN 881 110
FT DOMAIN 882 110
FT DOMAIN 883 110
FT DOMAIN 884 110
FT DOMAIN 885 110
FT DOMAIN 886 110
FT DOMAIN 887 110
FT DOMAIN 888 110
FT DOMAIN 889 110
FT DOMAIN 890 110
FT DOMAIN 891 110
FT DOMAIN 892 110
FT DOMAIN 893 110
FT DOMAIN 894 110
FT DOMAIN 895 110
FT DOMAIN 896 110
FT DOMAIN 897 110
FT DOMAIN 898 110
FT DOMAIN 899 110
FT DOMAIN 900 110
FT DOMAIN 901 110
FT DOMAIN 902 110
FT DOMAIN 903 110
FT DOMAIN 904 110
FT DOMAIN 905 110
FT DOMAIN 906 110
FT DOMAIN 907 110
FT DOMAIN 908 110
FT DOMAIN 909 110
FT DOMAIN 910 110
FT DOMAIN 911 110
FT DOMAIN 912 110
FT DOMAIN 913 110
FT DOMAIN 914 110
FT DOMAIN 915 110
FT DOMAIN 916 110
FT DOMAIN 917 110
FT DOMAIN 918 110
FT DOMAIN 919 110
FT DOMAIN 920 110
FT DOMAIN 921 110
FT DOMAIN 922 110
FT DOMAIN 923 110
FT DOMAIN 924 110
FT DOMAIN 925 110
FT DOMAIN 926 110
FT DOMAIN 927 110
FT DOMAIN 928 110
FT DOMAIN 929 110
FT DOMAIN 930 110
FT DOMAIN 931 110
FT DOMAIN 932 110
FT DOMAIN 933 110
FT DOMAIN 934 110
FT DOMAIN 935 110
FT DOMAIN 936 110
FT DOMAIN 937 110
FT DOMAIN 938 110
FT DOMAIN 939 110
FT DOMAIN 940 110
FT DOMAIN 941 110
FT DOMAIN 942 110
FT DOMAIN 943 110
FT DOMAIN 944 110
FT DOMAIN 945 110
FT DOMAIN 946 110
FT DOMAIN 947 110
FT DOMAIN 948 110
FT DOMAIN 949 110
FT DOMAIN 950 110
FT DOMAIN 951 110
FT DOMAIN 952 110
FT DOMAIN 953 110
FT DOMAIN 954 110
FT DOMAIN 955 110
FT DOMAIN 956 110
FT DOMAIN 957 110
FT DOMAIN 958 110
FT DOMAIN 959 110
FT DOMAIN 960 110
FT DOMAIN 961 110
FT DOMAIN 962 110
FT DOMAIN 963 110
FT DOMAIN 964 110
FT DOMAIN 965 110
FT DOMAIN 966 110
FT DOMAIN 967 110
FT DOMAIN 968 110
FT DOMAIN 969 110
FT DOMAIN 970 110
FT DOMAIN 971 110
FT DOMAIN 972 110
FT DOMAIN 973 110
FT DOMAIN 974 110
FT DOMAIN 975 110
FT DOMAIN 976 110
FT DOMAIN 977 110
FT DOMAIN 978 110
FT DOMAIN 979 110
FT DOMAIN 980 110
FT DOMAIN 981 110
FT DOMAIN 982 110
FT DOMAIN 983 110
FT DOMAIN 984 110
FT DOMAIN 985 110
FT DOMAIN 986 110
FT DOMAIN 987 110
FT DOMAIN 988 110
FT DOMAIN 989 110
FT DOMAIN 990 110
FT DOMAIN 991 110
FT DOMAIN 992 110
FT DOMAIN 993 110
FT DOMAIN 994 110
FT DOMAIN 995 110
FT DOMAIN 996 110
FT DOMAIN 997 110
FT DOMAIN 998 110
FT DOMAIN 999 110
FT DOMAIN 1000 110
FT DOMAIN 1001 110
FT DOMAIN 1002 110
FT DOMAIN 1003 110
FT DOMAIN 1004 110
FT DOMAIN 1005 110
FT DOMAIN 1006 110
FT DOMAIN 1007 110
FT DOMAIN 1008 110
FT DOMAIN 1009 110
FT DOMAIN 1010 110
FT DOMAIN
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

MM protein - protein search, using sw model

un on: November 21, 2003, 15:19:18 ; Search time 68.2397 Seconds  
(without alignments)  
286.100 Million cell updates/sec

File: US-09-010-377-2

Perfect score: 655

Sequence: 1 QVQLVQSGAEVKPKASVKV.....NYGVYMDYWGQGLTVTVSS 123

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	655	100.0	123	16 AAR81323	Humanized VLA-4 an
2	655	100.0	142	16 AAR81333	Human VLA-4 reshap
3	655	100.0	142	18 AAW22428	Humanised alpha-4
4	648	98.9	123	18 AAW22413	Humanised alpha-4
5	552	84.3	123	16 AAR81330	Mouse anti-VLA-4 a
6	544	83.1	140	16 AAR81327	Mouse VLA-4 antibo
7	544	83.1	140	18 AAW22410	Alpha-4 integrin m
8	523.5	79.9	120	19 AAW4124	Heavy chain variab
9	523.5	79.9	120	23 AAU79034	Humanised antibody

10	523.5	79.9	431	22 AAB30693	A fusion of single
11	523	79.8	136	16 AAR76681	Human/murine chime
12	523	79.8	136	17 AAW04396	Chimaeric human/mu
13	523	79.8	269	16 AAR76682	Human ONS-M21 anti
14	523	79.8	269	17 AAW04397	Chimaeric human/mu
15	515	78.6	119	16 AAR81325	Humanized VLA-4 an
16	515	78.6	119	18 AAW22426	Humanised alpha-4
17	512	78.2	119	16 AAR81324	Humanized VLA-4 an
18	512	78.2	119	18 AAW22425	Humanised alpha-4
19	507.5	77.5	135	21 AAB07969	A heavy chain vari
20	505.5	77.2	258	23 ABB05963	Monoclonal antibod
21	505.5	77.2	258	23 ABB05992	Mouse and human ch
22	505.5	77.2	258	23 ABB05996	Human monoclonal a
23	504.5	77.0	118	14 AAR37611	hIL2R Ab H chain v
24	503	76.8	119	20 AAY52717	Humanised ATR-5 H
25	503	76.8	119	22 AAB74978	Humanised ATR-5 H
26	501.5	76.6	258	23 ABB05991	Mouse and human ch
27	501.5	76.6	258	23 ABB05995	Human monoclonal a
28	499	76.2	119	20 AAY52719	Humanised ATR-5 H
29	499	76.2	119	22 AAB74980	Humanised ATR-5 H
30	496	75.7	136	17 AAB92084	Humanised antibody
31	496	75.7	136	23 ABB31443	Amino acid sequenc
32	495	75.6	119	16 AAR81331	Human 2*CL antibod
33	493	75.3	119	20 AAY52718	Humanised ATR-5 H
34	493	75.3	119	22 AAB74979	Humanised ATR-5 H
35	492	75.1	117	15 AAR57476	CDR-grafted anti-R
36	492	75.1	117	17 AAR32079	Marine 1308F VH CD
37	492	75.1	117	23 ABB31446	Humanised CDR-graf
38	492	75.1	136	15 AAR57481	Humanized 1308F VH
39	491.5	75.0	245	23 ABB45885	Human BlyS binding
40	491	75.0	119	20 AAY52708	Humanised ATR-5 H
41	491	75.0	119	22 AAB74959	Humanised ATR-5 H
42	489.5	74.7	120	18 AAW27551	Human Ab heavy cha
43	489	74.7	119	20 AAY52720	Humanised ATR-5 H
44	489	74.7	119	22 AAB74981	Humanised ATR-5 H
45	487.5	74.4	120	15 AAB74491	Humanised anti-CD1

ALIGNMENTS

RESULT 1

AAR81323

ID AAR81323 standard; Protein; 123 AA.

AC AAR81323;

DT 02-APR-1996 (first entry)

DE Humanized VLA-4 antibody 21.6 heavy chain variable region, Ha.

XX Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;

XX antibody engineering.

OS Mus musculus.

PN WO9519790-A1.

XX 27-JUL-1995.

PF 25-JAN-1995; 95WO-US01219.

PR 25-JAN-1994; 94US-0186269.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Bendig MM, Jones TS, Leger OJ, Saldanha J;

XX WPI; 1995-269276/35.

XX New humanised antibodies against VLA-4 - used for inhibiting

XX leucocyte adhesion to endothelial cells, partic. for treating

XX inflammatory disease.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: November 21, 2003, 15:21:54 ; Search time 23.1678 Seconds  
(without alignments)  
224.632 Million cell updates/sec

title: US-09-010-377-2

erfect score: 655

sequence: 1 QVQLVQSGAEVKPKGASVKY.....NYGVYANDYWGQGLTVVSS 123

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 328717 seqs, 42310858 residues

total number of hits satisfying chosen parameters: 328717

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- atabase :
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		Match	Length	DB ID	Description
	Score	%				
1	655	100.0	123	2	US-08-561-521-11	Sequence 11, Appl
2	655	100.0	123	5	PCT-US95-01219-11	Sequence 11, Appl
3	655	100.0	142	2	US-08-561-521-17	Sequence 17, Appl
4	655	100.0	142	5	PCT-US95-01219-17	Sequence 17, Appl
5	552	84.3	123	2	US-08-561-521-9	Sequence 9, Appl
6	552	84.3	123	5	PCT-US95-01219-9	Sequence 9, Appl
7	544	83.1	140	2	US-08-561-521-4	Sequence 4, Appl
8	544	83.1	140	5	PCT-US95-01219-4	Sequence 4, Appl
9	523.5	79.9	120	4	US-08-871-488A-19	Sequence 19, Appl
10	523	79.8	117	3	US-08-646-285A-132	Sequence 132, App
11	523	79.8	136	3	US-08-646-285A-99	Sequence 99, Appl
12	523	79.8	269	3	US-08-646-285A-109	Sequence 109, App
13	515	78.6	119	2	US-08-561-521-13	Sequence 13, Appl
14	515	78.6	119	5	PCT-US95-01219-13	Sequence 13, Appl
15	512	78.2	119	2	US-08-561-521-12	Sequence 12, Appl
16	512	78.2	119	5	PCT-US95-01219-12	Sequence 12, Appl
17	505	77.1	125	2	US-08-561-521-44	Sequence 44, Appl
18	505	77.1	125	5	PCT-US95-01219-44	Sequence 44, Appl
19	495	75.6	119	2	US-08-561-521-10	Sequence 10, Appl
20	495	75.6	119	5	PCT-US95-01219-10	Sequence 10, Appl
21	492	75.1	117	2	US-08-290-592E-17	Sequence 17, Appl
22	492	75.1	117	5	PCT-US95-10053-14	Sequence 14, Appl
23	492	75.1	117	5	PCT-US96-09448-17	Sequence 17, Appl
24	492	75.1	119	4	US-09-438-954-41	Sequence 41, Appl
25	491.5	75.0	118	2	US-08-232-081B-8	Sequence 8, Appl
26	489.5	74.7	120	4	US-09-025-769B-36	Sequence 36, Appl
27	489.5	74.7	120	4	US-09-025-769B-59	Sequence 59, Appl

28	486	74.2	140	3	US-08-836-561-63	Sequence 63, Appl
29	486	74.2	140	4	US-09-434-122-63	Sequence 63, Appl
30	484.5	74.0	135	1	US-08-137-117D-112	Sequence 112, App
31	484.5	74.0	135	2	US-08-436-717-112	Sequence 112, App
32	484.5	74.0	139	1	US-08-253-877C-19	Sequence 19, Appl
33	484.5	74.0	139	2	US-08-452-164A-19	Sequence 19, Appl
34	484.5	74.0	139	3	US-08-603-024-18	Sequence 14, Appl
35	484.5	74.0	139	4	US-08-450-809-14	Sequence 23, Appl
36	481	73.4	121	1	US-08-202-047-23	Sequence 23, Appl
37	481	73.4	121	3	US-08-964-690-23	Sequence 16, Appl
38	480.5	73.4	124	4	US-09-301-593-16	Sequence 41, Appl
39	480.5	73.4	124	4	US-09-301-593-41	Sequence 4, Appl
40	475	72.5	119	4	US-09-438-954-4	Sequence 74, Appl
41	474	72.4	140	3	US-08-836-561-74	Sequence 74, Appl
42	474	72.4	140	4	US-09-434-122-74	Sequence 53, Appl
43	473	72.2	123	1	US-08-482-882-53	Sequence 53, Appl
44	473	72.2	123	2	US-08-483-389-53	Sequence 53, Appl
45	473	72.2	123	2	US-08-487-113D-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1  
US-08-561-521-11  
; Sequence 11, Application US/08561521  
; Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; TITLE OF INVENTION: Adhesion Molecule VLA-4  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-561-521-11

Query Match 100.0%; Score 655; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2,7e-58;  
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 60  
| | | | |  
Db 1 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 60  
| | | | |

QY 61 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 120  
| | | | |  
Db 61 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 120  
| | | | |

QY 121 VSS 123  
| | |  
Db 121 VSS 123  
| | |

## RESULT 2

PCT-US95-01219-11  
; Sequence 11, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; TITLE OF INVENTION: Adhesion Molecule VIA-4  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01219  
; FILING DATE: 25-JAN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/186,269  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-01219-11

Query Match 100.0%; Score 655; DB 5; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2.7e-58;  
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 60  
| | | | |  
Db 1 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 60  
| | | | |

QY 61 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 120  
| | | | |  
Db 61 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 120  
| | | | |

QY 121 VSS 123  
| | |

Db 121 VSS 123  
| | |

## RESULT 3

US-08-561-521-17  
; Sequence 17, Application US/08561521  
; Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; TITLE OF INVENTION: Adhesion Molecule VIA-4  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 142 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-561-521-17

Query Match 100.0%; Score 655; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 3.2e-58;  
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 60  
| | | | |  
Db 20 QVQLVQSGAEVKKPGASVKVCKASGFIKDTYIHWRQAPGORLEWNGRIDPANGYTKY 79  
| | | | |

QY 61 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 120  
| | | | |  
Db 80 DPKFQGRVTITADTSASTAYMELSLRSEDVAVYCARREGYGNVGYVAMDYWGQGLT 139  
| | | | |

QY 121 VSS 123  
| | |  
Db 140 VSS 142  
| | |

## RESULT 4

PCT-US95-01219-17  
; Sequence 17, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

MM protein - protein search, using sw model

Run on: November 21, 2003, 15:27:54 ; Search time 45.9144 Seconds  
(without alignments)  
489.058 Million cell updates/sec

Title: US-09-010-377-2  
Perfect score: 655  
Sequence: 1 QVQLVQSGAEVKPEASVKV.....NYGVYMDYWGQTLVTVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.5	79.9	120	15	US-10-056-794-19
2	523.5	79.9	431	12	US-10-244-821-4
3	523.5	79.9	431	15	US-10-013-173-4
4	523.5	79.9	431	15	US-10-150-762-4
5	523	79.8	117	11	US-09-749-873-132
6	523	79.8	136	11	US-09-749-873-99
7	523	79.8	269	11	US-09-749-873-109
8	492	75.1	117	10	US-09-158-120A-17
9	492	75.1	119	15	US-10-233-996-41
10	491.5	75.0	245	11	US-09-880-748-1896
11	486	74.2	140	15	US-10-283-349-63
12	486	74.2	143	12	US-10-428-662-69
13	485.5	74.1	120	14	US-10-025-687-2
14	485.5	74.1	120	15	US-10-125-687-2
15	484	73.9	180	9	US-09-748-960-6
					Sequence 19, Appl
					Sequence 4, Appl
					Sequence 4, Appl
					Sequence 132, Appl
					Sequence 99, Appl
					Sequence 109, Appl
					Sequence 17, Appl
					Sequence 41, Appl
					Sequence 1896, Ap
					Sequence 63, Appl
					Sequence 69, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 6, Appl

16	480.5	73.4	124	12	US-10-159-006-16	Sequence 16, Appl
17	480.5	73.4	124	12	US-10-159-006-41	Sequence 41, Appl
18	480.5	73.4	124	15	US-10-121-464-16	Sequence 16, Appl
19	479.5	73.2	252	11	US-09-880-748-1698	Sequence 1698, Ap
20	477	72.8	141	11	US-09-791-153A-60	Sequence 60, Appl
21	476.5	72.7	249	11	US-09-880-748-1290	Sequence 1290, Ap
22	476.5	72.7	249	11	US-09-880-748-1299	Sequence 1299, Ap
23	476.5	72.7	251	11	US-09-880-748-1146	Sequence 1146, Ap
24	475	72.5	119	15	US-10-233-996-4	Sequence 4, Appl
25	474.5	72.4	255	11	US-09-880-748-1156	Sequence 1156, Ap
26	474	72.4	140	15	US-10-283-349-74	Sequence 74, Appl
27	473	72.2	123	9	US-09-753-436-53	Sequence 53, Appl
28	473	72.2	123	12	US-10-163-942-53	Sequence 53, Appl
29	473	72.2	123	15	US-10-269-805-27	Sequence 27, Appl
30	472.5	72.1	124	12	US-10-159-006-12	Sequence 12, Appl
31	472.5	72.1	124	12	US-10-159-006-40	Sequence 40, Appl
32	472.5	72.1	124	15	US-10-121-464-12	Sequence 12, Appl
33	472	72.1	123	9	US-09-753-436-86	Sequence 86, Appl
34	472	72.1	123	12	US-10-163-942-86	Sequence 86, Appl
35	472	72.1	136	15	US-10-160-232-91	Sequence 91, Appl
36	470	71.8	136	11	US-09-749-873-29	Sequence 29, Appl
37	469.5	71.7	126	11	US-09-791-153A-61	Sequence 61, Appl
38	469.5	71.7	230	11	US-09-791-153A-53	Sequence 53, Appl
39	469	71.6	129	10	US-09-956-206A-77	Sequence 77, Appl
40	469	71.6	470	10	US-09-859-053-28	Sequence 28, Appl
41	468.5	71.5	124	12	US-10-159-006-14	Sequence 14, Appl
42	468.5	71.5	124	12	US-10-159-006-108	Sequence 108, Appl
43	468.5	71.5	124	15	US-10-121-464-14	Sequence 14, Appl
44	468.5	71.5	249	11	US-09-880-748-1730	Sequence 1730, Ap
45	468	71.5	248	11	US-09-880-748-1178	Sequence 1178, Ap

ALIGNMENTS

RESULT 1  
US-10-056-794-19  
; Sequence 19, Application US/10056794  
; Publication No. US20030119078A1  
; GENERAL INFORMATION:  
; APPLICANT: Graves, Scott S.  
; Reno, John M.  
; Mallett, Robert W.  
; Hylarides, Mark D.  
; Searle, Stephen M.J.  
; Henry, Andrew H.  
; Pedersen, Jan T.  
; Rees, Anthony R.  
; TITLE OF INVENTION: HUMANIZED ANTIBODIES THAT BIND TO THE  
; ANTIGEN BOUND BY ANTIBODY NR-LU-13 AND THEIR USE IN  
; PRETARGETING METHODS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/056,794  
; FILING DATE: 24-Jan-2002  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey Ph.D., Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 690022.527C2  
; TELECOMMUNICATION INFORMATION:  
;

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-056-794-19

Query Match 79.9%; Score 523.5; DB 15; Length 120;  
Best Local Similarity 82.9%; Pred. No. 4.6e-43;  
Matches 102; Conservative 8; Mismatches 10; Indels 3; Gaps 1;  
QY 1 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKY 60  
Db 1 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKC 60  
QY 61 DPKEQGRVTITADTSASTAYMELSSLRSEDTAVVYCAREGYGNYGVYAMDYWGQGLT 120  
Db 61 DLSFGQGRVTITADTSINTAYMELSSLRSDDTAVVYCSREVL---TGTWSLDYWGQGLT 117  
QY 121 VSS 123  
Db 118 VSS 120

## RESULT 2

US-10-244-821-4  
Sequence 4, Application US/10244821  
Publication No. US20030143233A1  
GENERAL INFORMATION:  
APPLICANT: Goshorn, Stephen Charles  
APPLICANT: Graves, Scott Stoll  
APPLICANT: Schultz, Joanne Elaine  
APPLICANT: Lin, Yukang  
APPLICANT: Sanderson, James Allen  
APPLICANT: Reno, John M.  
APPLICANT: Dearstynne, Erica A.  
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: 690022.547C3  
CURRENT APPLICATION NUMBER: US/10/244,821  
CURRENT FILING DATE: 2002-09-16  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Predicted amino acid sequence for huNR-LU-10  
OTHER INFORMATION: single chain antibody-genomic streptavidin fusion.  
US-10-244-821-4

Query Match 79.9%; Score 523.5; DB 12; Length 431;  
Best Local Similarity 82.9%; Pred. No. 1.9e-42;  
Matches 102; Conservative 8; Mismatches 10; Indels 3; Gaps 1;  
QY 1 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKY 60  
Db 148 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKS 207  
QY 61 DPKEQGRVTITADTSASTAYMELSSLRSEDTAVVYCAREGYGNYGVYAMDYWGQGLT 120  
Db 208 DLSFGQGRVTITADTSINTAYMELSSLRSDDTAVVYCSREVL---TGTWSLDYWGQGLT 264  
QY 121 VSS 123  
Db 265 VSS 267

## RESULT 3

US-10-013-173-4  
Sequence 4, Application US/10013173  
Publication No. US20030095977A1  
GENERAL INFORMATION:  
APPLICANT: Goshorn, Stephen C.  
APPLICANT: Graves, Scott Stoll  
APPLICANT: Schultz, Joanne Elaine  
APPLICANT: Lin, Yukang  
APPLICANT: Sanderson, James A.  
APPLICANT: Reno, John M.  
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: 690022.547C1  
CURRENT APPLICATION NUMBER: US/10/013,173  
CURRENT FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Predicted amino acid sequence for huNR-LU-10  
OTHER INFORMATION: single chain antibody-genomic streptavidin fusion.  
US-10-013-173-4

Query Match 79.9%; Score 523.5; DB 15; Length 431;  
Best Local Similarity 82.9%; Pred. No. 1.9e-42;  
Matches 102; Conservative 8; Mismatches 10; Indels 3; Gaps 1;  
QY 1 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKY 60  
Db 148 QVQLVQSGAEVKPGASVKVSKASGFNIKDTYIHWVRQAPGQRLWGMGRIDPANGYTKS 207  
QY 61 DPKEQGRVTITADTSASTAYMELSSLRSEDTAVVYCAREGYGNYGVYAMDYWGQGLT 120  
Db 208 DLSFGQGRVTITADTSINTAYMELSSLRSDDTAVVYCSREVL---TGTWSLDYWGQGLT 264  
QY 121 VSS 123  
Db 265 VSS 267

## RESULT 4

US-10-150-762-4  
Sequence 4, Application US/10150762  
Publication No. US20030103948A1  
GENERAL INFORMATION:  
APPLICANT: Goshorn, Stephen C.  
APPLICANT: Graves, Scott S.  
APPLICANT: Schultz, Joanne E.  
APPLICANT: Lin, Yukang  
APPLICANT: Sanderson, James A.  
APPLICANT: Reno, John M.  
APPLICANT: Dearstynne, Erica A.  
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: 690022.547C2  
CURRENT APPLICATION NUMBER: US/10/150,762  
CURRENT FILING DATE: 2002-05-17  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Predicted amino acid sequence for huNR-LU-10  
OTHER INFORMATION: single chain antibody-genomic streptavidin fusion.  
US-10-150-762-4

Query Match

79.9%; Score 523.5; DB 15; Length 431;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: November 21, 2003, 15:20:03 ; Search time 12.2158 Seconds  
(without alignments)  
473.510 Million cell updates/sec

Title: US-09-010-377-2

Perfect score: 655

Sequence: 1 QVQLVQSGAEVKPGASVKV.....NYGVAMDYWGQGLTVTVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	427.5	65.3	147	1 HV1C HUMAN	P01744 homo sapien
2	409	62.4	120	1 HV03 MOUSE	P01747 mus musculus
3	408.5	62.4	139	1 HV07 MOUSE	P01751 mus musculus
4	405	61.8	117	1 HV1B HUMAN	P01743 homo sapien
5	399	60.9	140	1 HV02 MOUSE	P01746 mus musculus
6	396	60.5	117	1 HV1G HUMAN	P23083 homo sapien
7	390	59.5	117	1 HV1A HUMAN	P01742 homo sapien
8	378	57.7	117	1 HV13 MOUSE	P01757 mus musculus
9	375.5	57.3	118	1 HV51 MOUSE	P06330 mus musculus
10	373	56.9	117	1 HV12 MOUSE	P01755 mus musculus
11	371.5	56.7	137	1 HV11 MOUSE	P01755 mus musculus
12	369	56.3	138	1 HV48 MOUSE	P03980 mus musculus
13	364.5	55.6	120	1 HV50 MOUSE	P06329 mus musculus
14	360	55.0	121	1 HV01 MOUSE	P01745 mus musculus
15	360	55.0	136	1 HV15 MOUSE	P01759 mus musculus
16	355	54.2	125	1 HV1F HUMAN	P06326 homo sapien
17	349.5	53.4	120	1 HV1H HUMAN	P80421 homo sapien
18	349	53.3	117	1 HV09 HUMAN	P01753 mus musculus
19	344.5	52.6	114	1 HV00 MOUSE	P01741 mus musculus
20	338	51.6	117	1 HV04 MOUSE	P01748 mus musculus
21	336	51.3	117	1 HV06 MOUSE	P01750 mus musculus
22	333.5	50.9	136	1 HV16 MOUSE	P01783 mus musculus
23	333	50.8	117	1 HV10 MOUSE	P01754 mus musculus
24	330.5	50.5	124	1 HV1E HUMAN	P01761 homo sapien
25	329	50.2	117	1 HV14 MOUSE	P01758 mus musculus
26	329	50.2	117	1 HV49 MOUSE	P06328 mus musculus
27	326	49.8	117	1 HV52 MOUSE	P06327 mus musculus
28	326	49.8	123	1 HV24 MOUSE	P01793 mus musculus
29	325.5	49.7	119	1 HV17 MOUSE	P01807 mus musculus
30	325.5	49.7	124	1 HV1D HUMAN	P01760 homo sapien
31	324.5	49.5	119	1 HV38 MOUSE	P01808 mus musculus
32	323.5	49.4	119	1 HV40 MOUSE	P01810 mus musculus
33	320.5	48.9	122	1 HV3G HUMAN	P01768 homo sapien

#### ALIGNMENTS

##### RESULT 1

```

HV1C_HUMAN
ID HV1C_HUMAN STANDARD; PRT; 147 AA.
AC P01744;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-I region ND precursor [Fragments].
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8306234; PubMed=6815656;
RA Kenten J.H., Moigaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA Bell L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
RT immunoglobulin epsilon chain expressed in a myeloma cell line.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
RN [2]
RP SEQUENCE OF 20-147.
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL (In) Bach M.K. (eds.);
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
CC -!- MISCELLANEOUS: THIS EPSILON CHAIN WAS ISOLATED FROM A MYELOMA
CC PROTEIN.
-!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSP; P01789; IMCP.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SMO0406; Igv; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 19
FT CHAIN 20 147
FT DOMAIN 20 131
FT MOD_RES 20 20
FT DISULFID 41 115
FT CONFLICT 21 21
FT CONFLICT 53 54
FT CONFLICT 67 68
FT CONFLICT 125 125
FT NON_TER 147 147
SQ SEQUENCE 147 AA; 16491 MW; 9489F972A5366C20 CRC64;
```

Query Match

65.3%; Score 427.5; DB 1; Length 147;

Best Local Similarity 64.1%; Pred No. 1.1e-36;

Matches 82; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

```

QY 1 QVQLVQSGAEVKKPGASVKVSCKASGPNFKDTYIHWRQAPQORLEWGRIDPANGYTKY 60
Db 20 QQLVQSGAEVKKPGASVRSCKASGYTIDSYIHWRQAPGHGLEWGVNPNNGGTYN 79
QY 61 DKPQGRVTITADTSASTAYMELSLRSSEDTAVYYCAR-----EGYNGYGVYANDYWGQ 115
Db 80 APRFQGRVTMTDASFSTAYMQLSLRSLSRDSAVFYCAKSDPFWSDYINFDYSYTLDDWGQ 139
QY 116 GLTVTVSS 123
Db 140 GTTVTVSS 147

RESULT 2
HV03 MOUSE
ID HV03 MOUSE STANDARD; PRT; 120 AA.
AC P01747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 36-65.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=83131846; PubMed=6186498;
RA Siekevitz M., Geffer M.L., Brodeur P., Riblet R.,
RA Marshak-Rothstein A.;
RT "The genetic basis of antibody production: the dominant anti-arsenate
RT idiotypic response of the strain A mouse.";
RL Eur. J. Immunol. 12:1023-1032(1982).
CC -1- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER
CC DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE. THE AUTHORS
CC CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J
CC SEGMENT, JH2.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSSP: P01789; IMCP.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 111 IG-LIKE.
FT NON_TER 120
SQ SEQUENCE 120 AA; 13307 MW; FF04B4A167B654AF CRC64;

Query Match 62.4%; Score 409; DB 1; Length 120;
Best Local Similarity 63.1%; Pred. No. 6.5e-35;
Matches 77; Conservative 18; Mismatches 25; Indels 2; Gaps 1;

QY 2 QVQLVQSGAEVKKPGASVKVSCKASGPNFKDTYIHWRQAPQORLEWGRIDPANGYTKYD 61
Db 1 VQLQQSGAEVLRAGSSVKMSCKASGYTFTSVGINWVRQRPQCGLEMTGYINPGNGYTKYN 60
QY 62 PKPQGRVTITADTSASTAYMELSLRSSEDTAVYYCAREGYNGYGVYANDYWGQGLTVTV 121
Db 61 EKPKGKTLTVDKSSSTAYMQLSLRSLSRDSAVFYCARSVYIG--GSYFDYWGQGLTLTV 118
QY 122 SS 123
Db 119 SS 120

RESULT 3
HV07 MOUSE
ID HV07 MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

```

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region B1-8/186-2 precursor.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THE B1-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC (NPB ANTIBODIES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00529; AAA38170.1; -.
DR PIR; A90809; MHMS18.
DR PDB; 1AGU; 27-MAY-98.
DR PDB; 1AGW; 15-JUL-98.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region; Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 139 IG HEAVY CHAIN V REGION B1-8/186-2.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 124 D SEGMENT.
FT DOMAIN 125 139 JH2 SEGMENT.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15419 MW; 1B57DD4FDC09F465 CRC64;

Query Match 62.4%; Score 408.5; DB 1; Length 139;
Best Local Similarity 62.6%; Pred. No. 8.6e-35;
Matches 77; Conservative 18; Mismatches 25; Indels 3; Gaps 1;

QY 1 QVQLVQSGAEVKKPGASVKVSCKASGPNFKDTYIHWRQAPQORLEWGRIDPANGYTKY 60
Db 20 QVQLQSGAEVLRAGSSVKMSCKASGYTFTSVYINWVRQRPQCGLEWIGRIDPNSGGTKY 79
QY 61 DKPQGRVTITADTSASTAYMELSLRSSEDTAVYYCAREGYNGYGVYANDYWGQGLTVT 120
Db 80 NEKFKGKTLTVDKPSTAYMQLSLRSLSRDSAVFYCARVDYGS---SYFDYWGQGLTLT 136
QY 121 VSS 123
Db 137 VSS 139

RESULT 4
HV1B HUMAN
ID HV1B HUMAN STANDARD; PRT; 117 AA.
AC P01743;
DT 21-JUL-1986 (Rel. 01, Created)

```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	457.5	63.8	124	4	Q9UL92	Q9UL92 homo sapien
2	440.5	67.3	159	4	Q96Q80	Q96Q80 homo sapien
3	440	67.2	468	11	Q9UL31	Q9UL31 mus musculus
4	432	66.0	109	11	Q9UL85	Q9UL85 mus musculus
5	432	66.0	119	4	Q9UL94	Q9UL94 homo sapien
6	432	66.0	614	4	Q96G86	Q96G86 homo sapien
7	428	65.3	119	5	Q96Y22	Q96Y22 schistosoma
8	428	65.3	125	4	Q9UL95	Q9UL95 homo sapien
9	427	65.2	146	11	Q924Q3	Q924Q3 mus musculus
10	422	64.4	473	11	Q9D8L4	Q9D8L4 mus musculus
11	415.5	63.4	145	11	Q924R1	Q924R1 mus musculus
12	415.5	63.4	145	11	Q924R4	Q924R4 mus musculus
13	414.5	63.3	482	11	Q8K172	Q8K172 mus musculus
14	412	62.9	146	11	Q9UL88	Q9UL88 homo sapien
15	411.5	62.8	116	4	Q9UL89	Q9UL89 mus musculus
16	409.5	62.5	145	11	Q924Q6	Q924Q6 mus musculus

```
Query Match      69.8%; Score 457.5; DB 4; Length 124;  
Best Local Similarity 69.5%; Pred.No. 7.3e-41;  
Matches 91; Conservative 9; Mismatches 16; Indels 15; Gaps 2;
```

```
Db 1 EVOLVESGAEVKKPGASVKVSKASGYTFSSYNNHWVRQAPQGQLEWMGRIINPSGGSTSY 60
Qy 61 DPKFGQGVTTTADTSASTAYMELSSLRSEDYAVYVCAREGYNGYGVYAM-----DY 112
Db 61 AQKFGQGVTTTADTSASTAYMELSSLRSEDYAVYVCAR-----GLYVVVPAAFSFFDY 113
Qy 113 WQGGTLTVVSS 123
Db 114 WQGGTLTVVSS 124

RESULT 2
Q96QSO PRELIMINARY; PRT; 159 AA.
ID AC Q96QSO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative matrix cell adhesion molecule-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tilson M.D.;
RT "Homo sapiens putative microfibrillar protein with Ig-like domain 3
RL mRNA (Matrix Cell Adhesion Molecule-3, Mat-CAM 3).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY039025; AAK82649.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
SQ SEQUENCE 159 AA; 17497 MW; 5D29337B981FAF02 CRC64;

Query Match 67.3%; Score 440.5; DB 4; Length 159;
Best Local Similarity 64.9%; Pred. No. 6.4e-39;
Matches 87; Conservative 13; Mismatches 19; Indels 15; Gaps 2;

Qy 1 QVQLVQSGAEVKKPGASVKVSCKASGFIKDTYIHVWRQAPGQRLWNGRIDPANGYTKY 60
Db 20 QVQLVQSGAEVKKPGASVKVSCKASGYTFSSYNNHWVRQAPGQPEWNGVNPSSGSARY 79
Qy 61 DPKFGQGVTTTADTSASTAYMELSSLRSEDYAVYVCARE-----GYNGYGVYA 109
Db 80 SQKFGQGVTTTADTSASTAYMELSSLRSDDTAVYFCAREMEITFGGAVSKGFY----YYG 135

Qy 110 MDYWGQGLTVVSS 123
Db 136 MDVWGQGLTVVSS 149

RESULT 3
Q99L31 PRELIMINARY; PRT; 468 AA.
ID AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1; -.
DR HSP; F01842; 7FAB.
```

```
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B332ADB CRC64;

Query Match 67.2%; Score 440; DB 11; Length 468;
Best Local Similarity 67.5%; Pred. No. 2.8e-38;
Matches 83; Conservative 16; Mismatches 20; Indels 4; Gaps 1;

Qy 1 QVQLVQSGAEVKKPGASVKVSCKASGFIKDTYIHVWRQAPGQRLWNGRIDPANGYTKY 60
Db 20 EVQLQSGAEVLRPGASVKLSCTASGFIKDSLHWVKQRPQGLEWIGWIDPEDGETKY 79
Qy 61 DPKFGQGVTTTADTSASTAYMELSSLRSEDYAVYVCAREGYNGYGVYAMDYWGQGLTVT 120
Db 80 APKFDKATITADTSSTNTAYLQLSSLTSEDYAVYCARNLGYG----YDYWGQGLTIT 135
Qy 121 VSS 123
Db 136 VSS 138

RESULT 4
Q9JL85 PRELIMINARY; PRT; 109 AA.
ID AC Q9JL85;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anti-myosin immunoglobulin heavy chain variable region
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=20448942; PubMed=10992488;
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
RT with cardiac myosin."
RL Infect. Immun. 68:5803-5808 (2000).
DR EMBL; AF206021; AAP69319.1; -.
DR HSP; F01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 11944 MW; DFE615FE6CED4EDE CRC64;

Query Match 66.0%; Score 432; DB 11; Length 109;
Best Local Similarity 68.7%; Pred. No. 3.2e-38;
Matches 79; Conservative 15; Mismatches 15; Indels 6; Gaps 1;

Qy 9 AEVKKPGASVKVSCKASGFIKDTYIHVWRQAPGQRLWNGRIDPANGYTKYDPKFGQGV 68
Db 1 AELVKPGASVKLSCTASGFIKDTYMHVVKQRPQGLEWIGRIDPATGHSKYDKFGKA 60
Qy 69 TITADTSASTAYMELSSLRSEDYAVYVCAREGYNGYGVYAMDYWGQGLTVVSS 123
Db 61 TITSDTSNTAYLQLSSLTSEDYAVYCVRR-----GAVVFDYWGQGLTALTVSS 109
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: November 21, 2003, 15:21:54 , Search time 2.82534 Seconds  
(without alignments)  
224.632 Million cell updates/sec

itle: US-09-010-377-3

erfect score: 15  
equence: 1 XXXXXXXXXXXXX 15

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCITUS.COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		Query Match	Length	DB ID	Description
		Score	Match				
1	0	0.0	0.0	1	1	1	US-07-820-154A-12
2	0	0.0	0.0	1	1	1	US-07-791-213D-24
3	0	0.0	0.0	1	1	1	US-07-791-213D-40
4	0	0.0	0.0	1	1	1	US-08-174-365A-57
5	0	0.0	0.0	1	1	1	US-07-789-913-23
6	0	0.0	0.0	1	1	1	US-07-789-913-25
7	0	0.0	0.0	1	1	1	US-08-049-794-23
8	0	0.0	0.0	1	1	1	US-08-049-794-25
9	0	0.0	0.0	1	1	1	US-08-433-037-12
10	0	0.0	0.0	1	1	1	US-08-448-606-4
11	0	0.0	0.0	1	1	1	US-07-869-933-16
12	0	0.0	0.0	1	1	1	US-08-293-150A-24
13	0	0.0	0.0	1	1	1	US-08-293-150A-40
14	0	0.0	0.0	1	1	1	US-08-496-847-23
15	0	0.0	0.0	1	1	1	US-08-496-847-25
16	0	0.0	0.0	1	1	1	US-08-742-774-23
17	0	0.0	0.0	1	2	2	US-08-742-774-25
18	0	0.0	0.0	1	2	2	US-08-675-354-23
19	0	0.0	0.0	1	2	2	US-08-675-354-25
20	0	0.0	0.0	1	2	2	US-08-097-554A-12
21	0	0.0	0.0	1	2	2	US-08-965-918-23
22	0	0.0	0.0	1	2	2	US-08-965-918-25
23	0	0.0	0.0	1	2	2	US-09-138-439-23
24	0	0.0	0.0	1	2	2	US-09-138-439-25
25	0	0.0	0.0	1	3	3	US-08-480-640A-12
26	0	0.0	0.0	1	3	3	US-08-613-400A-23
27	0	0.0	0.0	1	3	3	US-08-613-400A-25

28	0	0.0	1	3	US-08-801-092-10	Sequence 10, Appl
29	0	0.0	1	3	US-08-801-092-17	Sequence 17, Appl
30	0	0.0	1	3	US-08-801-092-24	Sequence 24, Appl
31	0	0.0	1	3	US-08-801-092-31	Sequence 31, Appl
32	0	0.0	1	3	US-08-801-092-38	Sequence 38, Appl
33	0	0.0	1	3	US-08-801-092-45	Sequence 45, Appl
34	0	0.0	1	3	US-09-298-017-23	Sequence 23, Appl
35	0	0.0	1	3	US-09-298-017-25	Sequence 25, Appl
36	0	0.0	1	3	US-08-295-802-12	Sequence 12, Appl
37	0	0.0	1	3	US-09-332-379A-23	Sequence 23, Appl
38	0	0.0	1	3	US-09-332-379A-25	Sequence 25, Appl
39	0	0.0	1	3	US-09-103-663-16	Sequence 16, Appl
40	0	0.0	1	3	US-08-488-237A-12	Sequence 12, Appl
41	0	0.0	1	3	US-09-117-927-5	Sequence 5, Appl
42	0	0.0	1	4	US-08-375-992A-12	Sequence 12, Appl
43	0	0.0	1	4	US-09-315-113-10	Sequence 10, Appl
44	0	0.0	1	4	US-09-315-113-17	Sequence 17, Appl
45	0	0.0	1	4	US-09-315-113-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-07-820-154A-12  
; Sequence 12, Application US/07820154A  
; Patent No. 5382425  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Junker M.S., David E  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07820.154A  
; FILING DATE: 19920113  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 V 1

RESULT 2

US-07-791-213D-24

/ Sequence 24, Application US/07791213D  
/ Patent No. 5409895  
/ GENERAL INFORMATION:  
/ APPLICANT: MORISHITA, Hideaki  
/ APPLICANT: KANAMORI, Toshinori  
/ APPLICANT: NOBUHARA, Masahiro  
/ TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
/ TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
/ TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
/ TITLE OF INVENTION: TREATING USING THE SAME  
/ NUMBER OF SEQUENCES: 108  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Burns, Doane, Swecker & Mathis  
/ STREET: P.O. Box 1404  
/ CITY: Alexandria  
/ STATE: Virginia  
/ COUNTRY: United States  
/ ZIP: 22313-1404  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/07791,213D  
/ FILING DATE: 13-NOV-1991  
/ CLASSIFICATION: 435  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: JP 2-306745  
/ FILING DATE: 13-NOV-1990  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Meuth, Donna M  
/ REGISTRATION NUMBER: 36,607  
/ REFERENCE/DOCKET NUMBER: 029650-032  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 836-6620  
/ TELEFAX: (703) 836-2021  
/ INFORMATION FOR SEQ ID NO: 24:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-07-791-213D-24

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 3  
US-07-791-213D-40  
/ Sequence 40, Application US/07791213D  
/ Patent No. 5409895  
/ GENERAL INFORMATION:  
/ APPLICANT: MORISHITA, Hideaki  
/ APPLICANT: KANAMORI, Toshinori  
/ APPLICANT: NOBUHARA, Masahiro  
/ TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
/ TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
/ TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
/ TITLE OF INVENTION: TREATING USING THE SAME  
/ NUMBER OF SEQUENCES: 108  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Burns, Doane, Swecker & Mathis  
/ STREET: P.O. Box 1404  
/ CITY: Alexandria  
/ STATE: Virginia  
/ COUNTRY: United States

/ ZIP: 22313-1404  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/07791,213D  
/ FILING DATE: 13-NOV-1991  
/ CLASSIFICATION: 435  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: JP 2-306745  
/ FILING DATE: 13-NOV-1990  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Meuth, Donna M  
/ REGISTRATION NUMBER: 36,607  
/ REFERENCE/DOCKET NUMBER: 029650-032  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 836-6620  
/ TELEFAX: (703) 836-2021  
/ INFORMATION FOR SEQ ID NO: 40:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 4  
US-08-174-365A-57  
/ Sequence 57, Application US/08174365A  
/ Patent No. 5478809  
/ GENERAL INFORMATION:  
/ APPLICANT: Seiichi TANIDA et al  
/ TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES  
/ NUMBER OF SEQUENCES: 106  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Wenderoth, Lind & Ponack  
/ STREET: 805 Fifteenth Street, N.W., #700  
/ CITY: Washington  
/ STATE: D.C.  
/ COUNTRY: U.S.A.  
/ ZIP: 20005  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette, 3.5 inch, 144 mb  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: MS-DOS  
/ SOFTWARE: Wordperfect 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/174,365A  
/ FILING DATE: December 28, 1993  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER:  
/ FILING DATE:  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Warren M. Cheek, Jr.  
/ REGISTRATION NUMBER: 33,367  
/ REFERENCE/DOCKET NUMBER:  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 202-371-8850  
/ TELEFAX:  
/ TELEX:  
/ INFORMATION FOR SEQ ID NO: 57:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DOM protein - protein search, using sw model

Run on: November 21, 2003, 15:21:14 ; Search time 2.77397 Seconds  
(without alignments)  
520.023 Million cell updates/sec

Title: US-09-010-377-3

Perfect score: 15

Sequence: 1 XXXXXXXXXXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	0	0.0	3	RHTDIO	thyroliberin - Bom
2	0	0.0	3	RHPGT	thyroliberin - pig
3	0	0.0	3	RHSHT	thyroliberin - she
4	0	0.0	3	A92971	thyroliberin - eas
5	0	0.0	3	GKHU	growth-modulating
6	0	0.0	3	A60898	bursin - chicken
7	0	0.0	3	A23751	spinal cord peptid
8	0	0.0	3	B23751	spinal cord peptid
9	0	0.0	3	A33802	thyrotropin-releas
10	0	0.0	3	A22565	R-phycoerythrin al
11	0	0.0	3	PQ0010	angiotensin-conver
12	0	0.0	3	S13894	histidinol dehydro
13	0	0.0	3	A43391	TRH-like tripeptid
14	0	0.0	3	E37196	bradykinin-potentl
15	0	0.0	3	F37196	bradykinin-potentl
16	0	0.0	3	I50412	gene p20K protein
17	0	0.0	3	PT0636	T-cell receptor be
18	0	0.0	3	PT0578	T-cell receptor be
19	0	0.0	3	PT0571	T-cell receptor be
20	0	0.0	3	PT0622	T-cell receptor be
21	0	0.0	3	I78890	tyrosine protein k
22	0	0.0	3	S68328	blood cell protein
23	0	0.0	3	TL3892	cytochrome-c oxida
24	0	0.0	4	ECXAA	antho-RFamide neur
25	0	0.0	4	S18401	thyroglobulin - do
26	0	0.0	4	A02147	phagocytosis-stimu
27	0	0.0	4	A23039	tyrosine-melanocyt
28	0	0.0	4	ECNK	cardioexcitatory n
29	0	0.0	4	PL0140	carbon-monoxide de

30 0 0.0 4 2 PL0146 carbon-monoxide de  
31 0 0.0 4 2 A37832 phenol 2-monooxyge  
32 0 0.0 4 2 A48360 gamma subunit of p  
33 0 0.0 4 2 I40697 biotin A - Citroba  
34 0 0.0 4 2 A61300 22K superhelical D  
35 0 0.0 4 2 I57745 D-mannonate hydrol  
36 0 0.0 4 2 A41890 protein D - Escher  
37 0 0.0 4 2 S43014 hypothetical prote  
38 0 0.0 4 2 D41654 hypothetical prote  
39 0 0.0 4 2 B43848 cell surface adhes  
40 0 0.0 4 2 I40505 hypothetical prote  
41 0 0.0 4 2 I40870 phospholipase C (E  
42 0 0.0 4 2 I40804 endoglucanase F -  
43 0 0.0 4 2 T46627 starvation-induced  
44 0 0.0 4 2 S53508 ribosomal protein  
45 0 0.0 4 2 S17255

#### ALIGNMENTS

##### RESULT 1

RHTDIO

thyroliberin - Bombina orientalis

C:Species: Bombina orientalis

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A90919; A01415

R:Yasuhara, T.; Nakajima, T.

Chem. Pharm. Bull. 23, 3301-3303, 1975

A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.

A:Reference number: A90919; MUID:76138399; PMID:815011

A:Accession: A90919

A:Molecule type: protein

A:Residues: 1-3 <YAS>

C:Superfamily: thyroliberin precursor

F:1/Modified site: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Q 1

##### RESULT 2

RHPGT

thyroliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A01415

R:Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.

Biochemistry 9, 1103-1106, 1970

A:Title: Structure of porcine thyrotropin releasing hormone.

A:Reference number: A90560; MUID:70136150; PMID:4984938

A:Accession: A01415

A:Molecule type: protein

A:Residues: 1-3 <NAI>

R:Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.V.; Schally, A.V.

Biochem. Biophys. Res. Commun. 37, 705-710, 1969

A:Title: The identity of chemical and hormonal properties of the thyrotropin releasing hc

A:Reference number: A90167; MUID:70039904; PMID:4982117

A:Contents: annotation

A:Note: biological activities and Rf values (in 17 chromatographic systems) of the synthe

C:Superfamily: thyroliberin precursor

F:1/Modified site: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:3/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 3  
RHSHT  
thyroliberin - sheep  
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C/Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A93750; A01415  
R;Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.  
Org. Mass Spectrom. 5; 221-228, 1971  
A/Title: The elucidation of the primary structure of the hypothalamic thyroid stimulating hormone  
A/Reference number: A93750  
A/Accession: A93750  
A/Molecule type: protein  
A/Residues: 1-3 <DES>  
R;Burgess, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.  
Nature 226; 321-325, 1970  
A/Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.  
A/Reference number: A93161; MUID:70163386; PMID:4985794  
A/Contents: annotation  
A/Note: physicochemical characteristics and biological activities of the natural and synthetic thyroliberin precursor  
C/Superfamily: thyroliberin precursor  
C/Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F/1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F/3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 4  
A92971  
thyroliberin - eastern newt (tentative sequence)  
C/Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
C/Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A92971; A01415  
R;Grimm-Jorgensen, Y.; McKelvy, J.F.  
J. Neurochem. 23; 471-478, 1974  
A/Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain  
A/Reference number: A92971; MUID:75035605; PMID:4214528  
A/Accession: A92971  
A/Molecule type: protein  
A/Residues: 1-3 <GRI>  
A/Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid  
C/Superfamily: thyroliberin precursor  
C/Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic acid  
F/1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F/3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 5  
GXHU  
growth-modulating peptide - human

C/Species: Homo sapiens (man)  
C/Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A01421  
R;Schlesinger, D.H.; Pickart, L.; Thaler, M.M.  
Experientia 33; 324-325, 1977  
A/Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.  
A/Reference number: A01421; MUID:77162369; PMID:858356  
A/Accession: A01421  
A/Molecule type: protein  
A/Residues: 1-3 <SCH>  
A/Note: this serum tripeptide is found to stimulate growth of some cell types and to inhibit growth of others  
C/Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 G 1

RESULT 6  
A60898  
bursin - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A60898  
R;Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.  
Science 231; 997-999, 1986  
A/Title: tripeptide structure of bursin, a selective B-cell-differentiating hormone of the bursa of Fabricius  
A/Reference number: A60898; MUID:86122916; PMID:3484838  
A/Accession: A60898  
A/Molecule type: protein  
A/Residues: 1-3 <AUD>  
C/Superfamily: unassigned animal peptides  
C/Keywords: amidated carboxyl end; hormone  
F/3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 K 1

RESULT 7  
A23751  
spinal cord peptide SCP-4 - pig  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A23751  
R;Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240; 178-183, 1985  
A/Reference number: A23751; MUID:85250425; PMID:4015098  
A/Accession: A23751  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-3 <HSI>  
C/Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 8



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DOM protein - protein search, using sw model

Run on: November 21, 2003, 15:20:03 ; Search time 1.48973 Seconds  
(without alignments)  
473.510 Million cell updates/sec

Title: US-09-010-377-3  
Perfect score: 15  
Sequence: 1 XXXXXXXXXXXXXXX 15

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	1 GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	1 LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3	1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4	1 ACHI_ACHFU	P35904 achatina fu
5	0	0.0	4	1 DCML_PSECH	P19916 pseudomonas
6	0	0.0	4	1 DCMS_PSECH	P19918 pseudomonas
7	0	0.0	4	1 EOSI_HUMAN	P02731 homo sapien
8	0	0.0	4	1 PAR3_HIRME	P43562 hirudo medi
9	0	0.0	4	1 PAR4_HIRME	P43563 hirudo medi
10	0	0.0	4	1 FFKA_ATEL	P58705 anthopleura
11	0	0.0	4	1 FLRF_HIRME	P42561 hirudo medi
12	0	0.0	4	1 FLRN_ATEL	P58707 anthopleura
13	0	0.0	4	1 FMRP_MACNI	P01162 macrocallis
14	0	0.0	4	1 FVEI_ATEL	P58706 anthopleura
15	0	0.0	4	1 OCPI_OCTMI	P58648 octopus min
16	0	0.0	4	1 OCPI_OCTMI	P58649 octopus min
17	0	0.0	4	1 RM01_YEAST	P36515 saccharomyc
18	0	0.0	4	1 TUFT_HUMAN	P01858 homo sapien
19	0	0.0	5	1 ALI4_CARMA	P81817 carcinus ma
20	0	0.0	5	1 B10A_CITFR	P13071 citrobacter
21	0	0.0	5	1 B10B_CITFR	P12997 citrobacter
22	0	0.0	5	1 BPP7_BOTIN	P30425 bothrops in
23	0	0.0	5	1 EI03_LITRU	P82099 litoria rub
24	0	0.0	5	1 EI04_LITRU	P82100 litoria rub
25	0	0.0	5	1 FARP_ARTTR	P41853 artiopesthi
26	0	0.0	5	1 PAR2_PARMA	P81864 pardachirus
27	0	0.0	5	1 PRCT_PERAM	P01373 periplaneta
28	0	0.0	5	1 PSK_DAUCA	P58261 daucus caro
29	0	0.0	5	1 RE11_LITRU	P82070 litoria rub
30	0	0.0	5	1 RE21_LITRU	P82071 litoria rub
31	0	0.0	5	1 RE31_LITRU	P82072 litoria rub
32	0	0.0	5	1 RE32_LITRU	P82073 litoria rub
33	0	0.0	5	1 SUGA_ACHDO	P19991 acheta dome

34 0 0.0 5 1 TPIS\_CANFA P54714 canis faml  
35 0 0.0 5 1 TEM3\_ECOLI P13973 escherichia  
36 0 0.0 5 1 UC22\_MAIZE P80628 zea mays (m  
37 0 0.0 5 1 UF01\_MOUSE P38639 mus musculu  
38 0 0.0 5 1 UXA4\_CHLTR P38005 chlamydia t  
39 0 0.0 6 1 ACPH\_RABIT P25154 oryctolagus  
40 0 0.0 6 1 ASP2\_LACSN P82655 lactobacill  
41 0 0.0 6 1 CIP1\_MYTED P13736 mytilus edu  
42 0 0.0 6 1 CIP2\_MYTED P13737 mytilus edu  
43 0 0.0 6 1 EI01\_LITRU P82096 litoria rub  
44 0 0.0 6 1 FARP\_MONEK P41966 moniezia ex  
45 0 0.0 6 1 LOK1\_LOCM1 P41491 locusta mig

## ALIGNMENTS

RESULT 1  
GRWM\_HUMAN STANDARD; PRT; 3 AA.  
ID P01157;  
AC 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last annotation update)  
DE Growth-modulating peptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77162369; PubMed=858356;  
RA Schlesinger D.H., Pickart L., Thaler M.M.;  
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";  
RL Experimentia 33:324-325(1977).  
CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE  
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.  
DR GO; GO:0001558; P:regulation of cell growth; NAS.  
SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 X 1  
Db 1 G 1

RESULT 2  
LUXE\_VIBFI STANDARD; PRT; 3 AA.  
ID P24272;  
AC 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Long-chain-fatty-acid-luciferin-component ligase (EC 6.2.1.19) (Acyl-  
protein synthetase) (Fragment).  
GN LUXE.  
OS Vibrio fischeri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=668;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91072226; PubMed=2254256;  
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;  
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination  
RT site for the lux operon.";  
RL J. Bacteriol. 172:6797-6802(1990).  
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE  
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS

CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.  
 CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +  
 CC an acyl-L-protein thiolester.  
 CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M62812; -; NOT ANNOTATED\_CDS.  
 CC KW Luminescence; Ligase.  
 CC FT NON TER 1 1  
 CC SQ SEQUENCE 3 AA; 374 MW; 6AA3303000000000000 CRC64;  
 CC  
 CC Query Match 0.0%; Score 0; DB 1; Length 3;  
 CC Best Local Similarity 0.0%; Pred. No. 0;  
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 X 1  
 CC Db 1 I 1  
 CC  
 CC RESULT 3  
 CC THYL\_PIG  
 CC ID THYL\_PIG STANDARD; PRT; 3 AA.  
 CC AC P01151;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).  
 CC OS Sus scrofa (pig);  
 CC OS Ovis aries (Sheep);  
 CC OS Bombina orientalis (Oriental fire-bellied toad), and  
 CC OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).  
 CC OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 CC NCBI\_TaxID=9823, 9940, 8346, 8316;  
 CC [1]  
 CC SEQUENCE.  
 CC SPECIES=Pig; TISSUE=Hypothalamus;  
 CC MEDLINE=70136150; PubMed=4984938;  
 CC Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
 CC "Structure of porcine thyrotropin releasing hormone."  
 CC Biochemistry 9:1103-1106(1970).  
 CC [2]  
 CC SYNTHESIS.  
 CC SPECIES=Pig;  
 CC MEDLINE=70039904; PubMed=4982117;  
 CC Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
 CC "The identity of chemical and hormonal properties of the thyrotropin  
 CC releasing hormone and pyroglutamyl-histidyl-proline amide."  
 CC Biochem. Biophys. Res. Commun. 37:705-710(1969).  
 CC [3]  
 CC SEQUENCE.  
 CC SPECIES=Sheep; TISSUE=Hypothalamus;  
 CC Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,  
 CC Ward D.N.;  
 CC "The elucidation of the primary structure of the hypothalamic thyroid  
 CC stimulating hormone releasing factor of ovine origin by means of mass  
 CC spectrometry."  
 CC Org. Mass Spectrom. 5:221-228 (1971).  
 CC [4]  
 CC SYNTHESIS.  
 CC SPECIES=Sheep;  
 CC MEDLINE=70163386; PubMed=4985794;  
 CC Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,  
 CC Guillemin R.;  
 CC "Characterization of ovine hypothalamic hypophysiotropic

RT TSH-releasing factor.";  
 RL Nature 226:321-325 (1970).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=B.orientalis; TISSUE=Skin;  
 RX MEDLINE=76138399; PubMed=815011;  
 RA Yasuhara T., Nakajima T.;  
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin."  
 RL Chem. Pharm. Bull. 23:3301-3303(1975).  
 RN [6]  
 RP SEQUENCE.  
 RC SPECIES=N.viridescens;  
 RX MEDLINE=75035605; PubMed=4214528;  
 RA Grimm-Joergensen F., McKelvy J.F.;  
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus  
 viridescens) brain in vitro. Isolation and characterization of  
 thyrotropin releasing factor."  
 RL J. Neurochem. 23:471-478 (1974).  
 CC -!- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH  
 CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/  
 CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.  
 CC PIR; A90919; RHTDFO.  
 DR PIR; A92971; A92971.  
 DR PIR; A93750; RSHST.  
 KW Amidation; Pyrrolidone carboxylic acid.  
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD RES 3 3 AMIDATION.  
 SQ SEQUENCE 3 AA; 380 MW; 7761F6B0000000000000 CRC64;  
 CC  
 CC Query Match 0.0%; Score 0; DB 1; Length 3;  
 CC Best Local Similarity 0.0%; Pred. No. 0;  
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 X 1  
 CC Db 1 Q 1  
 CC  
 CC RESULT 4  
 CC ACHI\_ACHFU  
 CC ID ACHI\_ACHFU STANDARD; PRT; 4 AA.  
 CC AC P35954;  
 CC DT 01-JUN-1994 (Rel. 29, Created)  
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 CC DE Achatin-I.  
 CC OS Achatina fulica (Giant African snail).  
 CC OC Bukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 CC Sigmurethra; Achatinoidea; Achatinidae; Achatina.  
 CC NCBI\_TaxID=6530;  
 CC [1]  
 CC SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
 CC STRAIN=Ferussac; TISSUE=Ganglion;  
 CC MEDLINE=89273551; PubMed=2597281;  
 CC Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,  
 CC Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,  
 CC Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;  
 CC "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina  
 CC fulica Ferussac containing a D-amino acid residue."  
 CC Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
 CC [2]  
 CC CHARACTERIZATION.  
 CC STRAIN=Ferussac; TISSUE=Heart atrium;  
 CC MEDLINE=91264856; PubMed=1675568;  
 CC Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
 CC Yoshida M., Harada A., Muneoka Y., Kobayashi M.;  
 CC "Purification of achatin-I from the atria of the African giant snail,  
 CC Achatina fulica, and its possible function."  
 CC Biochem. Biophys. Res. Commun. 177:847-853(1991).  
 CC [3]  
 CC X-RAY CRYSTALLOGRAPHY.  
 CC MEDLINE=93014529; PubMed=1399265;  
 CC Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

MM protein - protein search, using sw model

Run on: November 21, 2003, 15:20:39 ; Search time 6.4726 Seconds  
(without alignments)  
598.027 Million cell updates/sec

Title: US-09-010-377-3

Perfect score: 15

Sequence: 1 XXXXXXXXXXXXXXX 15

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	0	0.0	4	11 Q08433	Q08433 rattus sp.
2	0	0.0	5	2 P83073	P83073 bacillus ce
3	0	0.0	5	10 Q99007	Q99007 hordeum vul
4	0	0.0	5	13 P83308	P83308 gallus gall
5	0	0.0	6	10 P82181	P82181 spinacia ol
6	0	0.0	6	10 P82541	P82541 spinacia ol
7	0	0.0	6	10 P82182	P82182 spinacia ol
8	0	0.0	7	2 Q8KMS3	Q8KMS3 klebsiella
9	0	0.0	7	2 Q47505	Q47505 escherichia
10	0	0.0	7	2 P70804	P70804 azotobacter
11	0	0.0	7	2 O50556	O50556 actinobacil
12	0	0.0	7	2 Q47477	Q47477 escherichia
13	0	0.0	7	2 Q47029	Q47029 enterobacte
14	0	0.0	7	2 P72081	P72081 nocardia la
15	0	0.0	7	2 Q54248	Q54248 streptomyc
16	0	0.0	7	2 Q8KMS9	Q8KMS9 enterobacte

17	0	0.0	7	2 Q34028	Q34028 sphingomona
18	0	0.0	7	2 Q07354	Q07354 synechococc
19	0	0.0	7	2 Q8GL12	Q8GL12 borrelia bu
20	0	0.0	7	2 Q8GL04	Q8GL04 borrelia bu
21	0	0.0	7	2 Q8GL00	Q8GL00 borrelia bu
22	0	0.0	7	4 Q15903	Q15903 homo sapien
23	0	0.0	7	4 Q8NH7	Q8NH7 homo sapien
24	0	0.0	7	4 Q15897	Q15897 homo sapien
25	0	0.0	7	6 Q28742	Q28742 oryctolagus
26	0	0.0	7	8 P92214	P92214 amblyopryzum
27	0	0.0	7	8 P92393	P92393 hordeum vul
28	0	0.0	7	8 P92403	P92403 lophopyrum
29	0	0.0	7	8 P92427	P92427 peridictyon
30	0	0.0	7	8 Q99182	Q99182 gnatholebia
31	0	0.0	7	8 P92430	P92430 aegilops ta
32	0	0.0	7	8 P92221	P92221 bromus iner
33	0	0.0	7	8 Q98866	Q98866 spinacia ol
34	0	0.0	7	8 P92425	P92425 pseudoroegn
35	0	0.0	7	8 P92381	P92381 hordeum bra
36	0	0.0	7	8 P92387	P92387 henrardia p
37	0	0.0	7	8 P92210	P92210 agropyron c
38	0	0.0	7	8 P92440	P92440 thinopyrum
39	0	0.0	7	8 P92218	P92218 australopyr
40	0	0.0	7	8 P92390	P92390 heteranthe
41	0	0.0	7	8 P92372	P92372 haynaldia v
42	0	0.0	7	8 P92442	P92442 taeniathe
43	0	0.0	7	8 P92226	P92226 crithopsis
44	0	0.0	7	8 Q8MPY6	Q8MPY6 taraxacum (
45	0	0.0	7	8 Q95945	Q95945 saccharomyC

#### ALIGNMENTS

##### RESULT 1

Q08433 PRELIMINARY; PRT; 4 AA.  
AC Q08433;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Bilirubin UDP-glucuronosyltransferase (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Gunn;  
RX MEDLINE=91282758; PubMed=1840486;  
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;  
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the  
hyperbilirubinemic Gunn rat.";  
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
DR EMBL; S38636; AAB19259.1; -.  
KW Transferase.  
FT NON TER 1  
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 N 1

##### RESULT 2

P83073 PRELIMINARY; PRT; 5 AA.  
ID P83073  
AC P83073;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE 88 kDa protein (Fragment).  
OS Bacillus cereus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1396;  
[1]  
RP SEQUENCE.  
RC STRAIN=NCIMB 11796;  
RA Brown N., Dows B.C.A.;  
RT Submitted (JUL-2001) to the SWISS-PROT data bank.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;  
[1]  
Query Match 0.0%; Score 0; DB 2; Length 5;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 4 T 4  
RESULT 3  
Q99007 PRELIMINARY; PRT; 5 AA.  
ID Q99007;  
AC Q99007;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Alpha amylase (Fragment).  
GN AMY1 GENE.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
[1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=91329704; PubMed=1831085;  
RA Jacobsen J.V., Close T.J.;  
RT "Control of transient expression of chimaeric genes by gibberellic acid and abscisic acid in protoplasts prepared from mature barley aleurone layers."  
RT Plant Mol. Biol. 16:713-721(1991).  
DR EMBL; X54643; CAA38455.1; -.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
Query Match 0.0%; Score 0; DB 10; Length 5;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 2 A 2  
RESULT 4  
P83308 PRELIMINARY; PRT; 5 AA.  
ID P83308;  
AC P83308;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DE FMRamide-like neuropeptide (IPURF-amide).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
[1]  
RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Brain;  
RX PubMed=6137771;  
RA Dockray G.J., Reeve J.R. Jr., Shively J., Cayton R.J., Barnard C.S.;  
RT "A novel active pentapeptide from chicken brain identified by antibodies to FMRamide."  
RL Nature 305:328-330(1983).  
CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE) FAMILY.  
KW Neuropeptide.  
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;  
Query Match 0.0%; Score 0; DB 13; Length 5;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 L 1  
RESULT 5  
P82181 PRELIMINARY; PRT; 6 AA.  
ID P82181;  
AC P82181;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
[1]  
RP SEQUENCE.  
RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
DR InterPro; IPR001790; Ribosomal\_L10.  
DR InterPro; IPR002363; Ribosomal\_L10eub.  
DR Pfam; PF00466; Ribosomal\_L10; PARTIAL.  
DR PROSITE; PS01109; RIBOSOMAL\_L10; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER 6  
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB0000 CRC64;  
Query Match 0.0%; Score 0; DB 10; Length 6;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 A 1  
RESULT 6  
P82541 PRELIMINARY; PRT; 6 AA.  
ID P82541;  
AC P82541;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
OS Spinacia oleracea (Spinach).

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

run on: November 21, 2003, 15:19:18 ; Search time 16.6438 Seconds  
(without alignments)  
286.100 Million cell updates/sec

us-09-010-377-4

Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Genes19Jun03.\*  
1: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.\*  
2: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.\*  
3: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.\*  
4: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.\*  
5: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.\*  
6: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.\*  
7: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.\*  
8: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.\*  
9: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.\*  
10: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.\*  
11: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.\*  
12: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.\*  
13: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.\*  
14: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.\*  
15: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.\*  
16: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.\*  
17: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.\*  
18: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.\*  
19: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.\*  
20: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.\*  
21: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.\*  
22: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.\*  
23: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.\*  
24: /SIDSI1/gcgdata/geneseq/geneseq-embl/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	0	0.0	1	20 AAY45652	Immunogenic peptid
2	0	0.0	1	22 ABB56870	Human SNP related
3	0	0.0	1	22 ABB68809	Drosophila melanog
4	0	0.0	1	22 ABB66810	Drosophila melanog
5	0	0.0	1	22 ABB02941	Novel human diagn
6	0	0.0	1	22 AAM97643	Human peptide #918
7	0	0.0	1	22 AAM97834	Human peptide #910
8	0	0.0	1	22 AAM97974	Human peptide #124
9	0	0.0	1	22 AAM98354	Human peptide #162

10	0	0.0	1	22 AAM98447	Human peptide #172
11	0	0.0	1	22 AAM53218	Human nonconservat
12	0	0.0	1	22 AAM53219	Human nonconservat
13	0	0.0	1	22 AAM53290	Human nonconservat
14	0	0.0	1	22 AAM53291	Human nonconservat
15	0	0.0	1	22 AAM53328	Human nonconservat
16	0	0.0	1	22 AAM53329	Human nonconservat
17	0	0.0	1	22 AAM45230	H11 binding site c
18	0	0.0	1	22 AAG99966	ERA binding domain
19	0	0.0	1	22 AAG99983	ERA binding domain
20	0	0.0	1	22 AAG99987	ERA binding domain
21	0	0.0	1	22 AAG99988	ERA binding domain
22	0	0.0	1	22 AAM00010	ERA binding domain
23	0	0.0	1	22 AAM00011	ERA binding domain
24	0	0.0	1	22 AAM00013	ERA binding domain
25	0	0.0	1	22 AAM00016	Human SNP associat
26	0	0.0	1	22 AAG98026	Human SNP associat
27	0	0.0	1	22 AAG98134	Thyrotropin releas
28	0	0.0	1	22 AAB91029	Thyrotropin releas
29	0	0.0	1	22 AAB91546	Endothelins and re
30	0	0.0	1	22 AAB91655	Opioid peptide SEQ
31	0	0.0	1	22 AAB91739	Opioid peptide SEQ
32	0	0.0	1	22 AAB91892	Apoptosis related
33	0	0.0	1	22 AAB92150	Polypeptide SEQ ID
34	0	0.0	1	22 AAB92392	Miscellaneous pept
35	0	0.0	1	23 ABUS1078	Helicobacter pylor
36	0	0.0	1	23 ABUS1919	Helicobacter pylor
37	0	0.0	1	23 ABUS2012	Helicobacter pylor
38	0	0.0	1	23 ABUS2118	Helicobacter pylor
39	0	0.0	1	23 ABG95255	Human novel secret
40	0	0.0	1	23 ABG95471	Human novel secret
41	0	0.0	1	23 ABG93483	Human P-glycoprote
42	0	0.0	1	23 ABG93484	Human P-glycoprote
43	0	0.0	1	23 ABG93489	Human P-glycoprote
44	0	0.0	1	23 ABG93491	Human P-glycoprote
45	0	0.0	1	23 ABG93499	Human P-glycoprote

ALIGNMENTS

RESULT 1  
AAY46652  
ID AAY46652 standard; Peptide; 1 AA.  
XX AC AAY46652;  
XX 01-DEC-1999 (first entry)  
XX Immunogenic peptide having a human leukocyte antigen binding motif #1263.  
DE Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
XX immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.  
XX Synthetic.  
OS Homo sapiens.  
XX PN WO9945954-A1.  
XX PD 16-SEP-1999.  
XX PF 13-MAR-1998; 98WO-US05039.  
XX PR 13-MAR-1998; 98WO-US05039.  
XX PA (EPIM-) EPIMUNE INC.  
XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
XX WPI; 1999-551214/46.  
DR

XX New immunogenic peptides with HLA binding motif, useful in treatment  
 PPT and diagnosis of cancers and viral diseases -  
 XX  
 XX  
 XX  
 XXX Claim 1; Page 80; 150pp; English.

AAV45390 to AAY48214 represent specifically claimed immunogenic peptides having a human major histocompatibility complex (MHC) Class I (also known as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against the antigen from which the peptide is derived. Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are normally induced by an antigen in the form of a peptide fragment bound to a HLA molecule, rather than the intact foreign antigen itself, and are particularly important in tumour rejection and in fighting viral infections. The peptides are therefore useful therapeutically to treat or prevent viral infections and cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma. They can be administered as vaccines to elicit an immune response in individuals susceptible or otherwise at risk of viral infection or cancer, or used to treat chronic or acute conditions. They are also useful diagnostically, and can be used to induce a cytotoxic T cell response by contacting a cytotoxic T cell with the peptide e.g. to produce CTLs ex vivo for infusion back into a patient. The peptide e.g. to polynucleotides encoding the immunogenic peptides are also useful therapeutically and for immunisation as above.

Sequence 1 AA;  
 Query Match 0.0%; Score 0; DB 20; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 Y 1

RESULT 2  
 ABB56870  
 IID ABB56870 standard; Peptide; 1 AA.  
 XX  
 XX ABB56870;  
 DT AC  
 DT XX  
 DT XX  
 DE XX  
 XX  
 XX Human SNP related amino acid sequence SEQ ID NO:1435.  
 KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic; immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;  
 KW autoimmune disease; inflammation; cancer; nervous system disease;  
 KW infection; polymorphic protein.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200138586-A2.  
 XX  
 XX  
 PD 31-MAY-2001.  
 XX  
 XX 22-NOV-2000; 2000WO-US23211.  
 XX  
 XX 24-NOV-1999; 99US-0167383.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach M;  
 XX WPI; 2001-355949/37.  
 XX  
 XX Isolated human nucleic acids comprising one or more single nucleotide  
 PPT polymorphisms, useful for treating a subject suffering from a  
 PPT pathology, e.g. autoimmune diseases, ascribed to the presence of a

```

PT sequence polymorphism -
XX Claim 1; Page 666; 674pp; English.
XX
XX ABL000010 to ABL01104 represent human nucleic acid oligonucleotides
XX comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
XX to ABB56903 represent human peptides encoded by some of the SNP
XX oligonucleotides. The sequences from the present invention can have
XX immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
XX antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
XX and antibodies from the present invention can be used for treating a
XX subject suffering from, at risk for, or suspected of, suffering from a
XX pathology ascribed to the presence of a sequence polymorphism. The
XX pathology may be autoimmune diseases, inflammation, cancer, diseases of
XX the nervous system, and infection by pathogenic microorganisms. The SNPs
XX are also useful for determining which forms of a characterised
XX polymorphism are present in individuals. The antibodies may be used in
XX the detection, quantitation and/or cellular or tissue localisation of a
XX polymorphic protein (e.g., for use in measuring levels of the
XX polymorphic protein within appropriate physiological samples).
XX
XX Sequence 1 AA;
XX
XX Query Match 0.0%; Score 0; DB 22; Length 1;
XX Best Local Similarity 0.0%; Pred. No. 0;
XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0
XX
XX QY 1 X 1
XX DB 1 S 1
XX
XX
XX RESULT 3
XX ABB66809
XX ID ABB66809 standard; Protein; 1 AA.
XX
XX AC ABB66809;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 27219.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX FR 11-JUL-2000; 2000US-0614150.
XX
XX (PEXE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL10912.
XX
XX
XX PT Genes isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX
XX PS Disclosure; SEQ ID NO 27219; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 15:21:54 ; Search time 5.65068 Seconds  
(without alignments)  
224.632 Million cell updates/sec

Title: US-09-010-377-4  
Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table: BLOSUMP62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	0	0.0	1	1	US-07-820-154A-12
2	0	0.0	1	1	US-07-791-213D-24
3	0	0.0	1	1	US-07-791-213D-40
4	0	0.0	1	1	US-08-174-365A-57
5	0	0.0	1	1	US-07-789-913-23
6	0	0.0	1	1	US-07-789-913-25
7	0	0.0	1	1	US-08-049-794-23
8	0	0.0	1	1	US-08-049-794-25
9	0	0.0	1	1	US-08-433-037-12
10	0	0.0	1	1	US-08-448-606-4
11	0	0.0	1	1	US-07-869-933-16
12	0	0.0	1	1	US-08-293-150A-24
13	0	0.0	1	1	US-08-293-150A-40
14	0	0.0	1	1	US-08-496-847-23
15	0	0.0	1	1	US-08-496-847-25
16	0	0.0	1	2	US-08-742-774-23
17	0	0.0	1	2	US-08-742-774-25
18	0	0.0	1	2	US-08-675-354-23
19	0	0.0	1	2	US-08-675-354-25
20	0	0.0	1	2	US-08-097-554A-12
21	0	0.0	1	2	US-08-965-918-23
22	0	0.0	1	2	US-08-965-918-25
23	0	0.0	1	2	US-09-138-439-23
24	0	0.0	1	2	US-09-138-439-25
25	0	0.0	1	3	US-08-480-640A-12
26	0	0.0	1	3	US-08-613-400A-23
27	0	0.0	1	3	US-08-613-400A-25

28	0	0.0	1	3	US-08-801-092-10	Sequence 10, Appl
29	0	0.0	1	3	US-08-801-092-17	Sequence 17, Appl
30	0	0.0	1	3	US-08-801-092-24	Sequence 24, Appl
31	0	0.0	1	3	US-08-801-092-31	Sequence 31, Appl
32	0	0.0	1	3	US-08-801-092-38	Sequence 38, Appl
33	0	0.0	1	3	US-08-801-092-45	Sequence 45, Appl
34	0	0.0	1	3	US-09-298-017-23	Sequence 23, Appl
35	0	0.0	1	3	US-09-298-017-25	Sequence 25, Appl
36	0	0.0	1	3	US-08-295-802-12	Sequence 12, Appl
37	0	0.0	1	3	US-09-392-979A-23	Sequence 23, Appl
38	0	0.0	1	3	US-09-392-979A-25	Sequence 25, Appl
39	0	0.0	1	3	US-09-103-663-16	Sequence 16, Appl
40	0	0.0	1	3	US-08-488-237A-12	Sequence 12, Appl
41	0	0.0	1	3	US-09-117-927-5	Sequence 5, Appl
42	0	0.0	1	4	US-08-375-992A-12	Sequence 12, Appl
43	0	0.0	1	4	US-09-315-113-10	Sequence 10, Appl
44	0	0.0	1	4	US-09-315-113-17	Sequence 17, Appl
45	0	0.0	1	4	US-09-315-113-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-07-820-154A-12  
; Sequence 12, Application US/07820154A  
; Patent No. 5382425  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Junker M.S., David E  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/820,154A  
; FILING DATE: 19920113  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)977-9550  
; TELEFAX: (212)664-0525  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 V 1

RESULT 2  
US-07-791-213D-24

; Sequence 24, Application US/07791213D  
; Patent No. 5409895  
; GENERAL INFORMATION:  
; APPLICANT: MORISHITA, Hideaki  
; APPLICANT: KANAMORI, Toshinori  
; APPLICANT: NOBUHARA, Masahiro  
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
; TITLE OF INVENTION: TREATING USING THE SAME  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/791,213D  
; FILING DATE: 13-NOV-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 13-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meuth, Donna M  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 029650-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-791-213D-24

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 C 1

RESULT 3  
US-07-791-213D-40  
; Sequence 40, Application US/07791213D  
; Patent No. 5409895  
; GENERAL INFORMATION:  
; APPLICANT: MORISHITA, Hideaki  
; APPLICANT: KANAMORI, Toshinori  
; APPLICANT: NOBUHARA, Masahiro  
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
; TITLE OF INVENTION: TREATING USING THE SAME  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States

; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/791,213D  
; FILING DATE: 13-NOV-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 13-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meuth, Donna M  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 029650-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 C 1

RESULT 4  
US-08-174-365A-57  
; Sequence 57, Application US/08174365A  
; Patent No. 5478809  
; GENERAL INFORMATION:  
; APPLICANT: Seiichi TANIDA et al.  
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/174,365A  
; FILING DATE: December 28, 1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 57:



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 15:27:54 ; Search time 11.1986 Seconds  
(without alignments)  
489.058 Million cell updates/sec

Title: US-09-010-377-4  
Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 18259486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	0	0.0	1	9	US-09-909-348-4
2	0	0.0	1	10	US-09-982-172-3
3	0	0.0	1	10	US-09-982-172-4
4	0	0.0	1	10	US-09-982-172-9
5	0	0.0	1	10	US-09-982-172-11
6	0	0.0	1	10	US-09-982-172-19
7	0	0.0	1	10	US-09-982-172-31
8	0	0.0	1	10	US-09-982-172-35
9	0	0.0	1	10	US-09-982-172-37
10	0	0.0	1	10	US-09-982-172-46
11	0	0.0	1	10	US-09-982-172-69
12	0	0.0	1	10	US-09-982-172-80
13	0	0.0	1	10	US-09-982-172-81
14	0	0.0	1	10	US-09-982-172-83
15	0	0.0	1	10	US-09-982-172-86

16	0	0.0	1	10	US-09-982-172-93	Sequence 93, Appl
17	0	0.0	1	10	US-09-982-172-95	Sequence 95, Appl
18	0	0.0	1	10	US-09-982-172-106	Sequence 106, Appl
19	0	0.0	1	10	US-09-982-172-112	Sequence 112, Appl
20	0	0.0	1	10	US-09-982-172-120	Sequence 120, Appl
21	0	0.0	1	10	US-09-982-172-126	Sequence 126, Appl
22	0	0.0	1	10	US-09-982-172-148	Sequence 148, Appl
23	0	0.0	1	10	US-09-982-172-149	Sequence 149, Appl
24	0	0.0	1	10	US-09-982-172-155	Sequence 155, Appl
25	0	0.0	1	10	US-09-982-172-160	Sequence 160, Appl
26	0	0.0	1	10	US-09-982-172-172	Sequence 172, Appl
27	0	0.0	1	10	US-09-982-172-173	Sequence 173, Appl
28	0	0.0	1	10	US-09-982-172-175	Sequence 175, Appl
29	0	0.0	1	10	US-09-982-172-189	Sequence 189, Appl
30	0	0.0	1	10	US-09-982-172-190	Sequence 190, Appl
31	0	0.0	1	10	US-09-982-172-191	Sequence 191, Appl
32	0	0.0	1	10	US-09-982-172-195	Sequence 195, Appl
33	0	0.0	1	10	US-09-982-172-200	Sequence 200, Appl
34	0	0.0	1	10	US-09-982-172-211	Sequence 211, Appl
35	0	0.0	1	11	US-09-809-391-395	Sequence 395, Appl
36	0	0.0	1	11	US-09-809-391-611	Sequence 611, Appl
37	0	0.0	1	12	US-10-032-221B-19	Sequence 19, Appl
38	0	0.0	1	12	US-10-299-619-4	Sequence 4, Appl
39	0	0.0	1	12	US-10-061-607A-28	Sequence 28, Appl
40	0	0.0	1	12	US-10-061-607A-29	Sequence 29, Appl
41	0	0.0	1	12	US-10-061-607A-35	Sequence 35, Appl
42	0	0.0	1	12	US-09-882-171-395	Sequence 395, Appl
43	0	0.0	1	12	US-09-882-171-611	Sequence 611, Appl
44	0	0.0	1	12	US-10-104-307-7	Sequence 7, Appl
45	0	0.0	1	12	US-10-104-307-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-909-348-4  
; Sequence 4, Application US/09909348  
; Patent No. US20020042373A1  
; GENERAL INFORMATION:  
; APPLICANT: Carney, Darrell H.  
; APPLICANT: Crowther, Roger S.  
; APPLICANT: Stierberg, Janet  
; APPLICANT: Bergmann, John  
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists  
; TITLE OF INVENTION: Of The No. US20020042373A1-Proteolytically Activated Thrombin Rec  
; FILE REFERENCE: 3033.1003-001  
; CURRENT APPLICATION NUMBER: US/09/909,348  
; CURRENT FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US 60/219,800  
; PRIOR FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide fragment of Thrombin  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(14)  
; OTHER INFORMATION: Xaa at position six is Glu or Gln  
; OTHER INFORMATION: Xaa at position thirteen is Phe, Met, Leu, His or  
US-09-909-348-4

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 V 1

```
RESULT 2
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3
Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 3
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4
Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 4
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9
Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 5
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11
Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 6
US-09-982-172-19
; Sequence 19, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-19
Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 15:21:14 ; Search time 5.54795 Seconds  
(without alignments)  
520.023 Million cell updates/sec

Title: US-09-010-377-4  
Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	0	0.0	3	3 RHTDIO	thyroliberin - Bom
2	0	0.0	3	3 RHPGT	thyroliberin - pig
3	0	0.0	3	3 RHSHT	thyroliberin - she
4	0	0.0	3	3 A92371	thyroliberin - eas
5	0	0.0	3	3 GKRU	growth-modulating
6	0	0.0	3	3 A60898	bursin - chicken
7	0	0.0	3	3 A23751	spinal cord peptid
8	0	0.0	3	3 B23751	spinal cord peptid
9	0	0.0	3	3 A33802	thyrotropin-releas
10	0	0.0	3	3 A22565	R-phycoerythrin al
11	0	0.0	3	3 PQ0010	angiotensin-conver
12	0	0.0	3	3 S13894	histidinol dehydro
13	0	0.0	3	3 A43391	TRH-like tripeptid
14	0	0.0	3	3 E37196	bradykinin-potentl
15	0	0.0	3	3 F37196	bradykinin-potentl
16	0	0.0	3	3 I50412	gene p20K protein
17	0	0.0	3	3 PT0636	T-cell receptor be
18	0	0.0	3	3 PT0578	T-cell receptor be
19	0	0.0	3	3 PT0571	T-cell receptor be
20	0	0.0	3	3 PT0622	tyrosine protein k
21	0	0.0	3	3 I78890	blood cell protein
22	0	0.0	3	3 S63228	cytochrome-c oxida
23	0	0.0	3	3 T13892	antho-rfamide neur
24	0	0.0	4	1 ECXAA	thyroglobulin - do
25	0	0.0	4	2 S18401	phagocytosis-stimu
26	0	0.0	4	2 A02147	tyrosine-melanocyt
27	0	0.0	4	2 A32039	cardioexcitatory n
28	0	0.0	4	2 ECNK	carbon-monoxide de
29	0	0.0	4	2 PL0140	

30	0	0.0	4	2 PL0146	carbon-monoxide de
31	0	0.0	4	2 A37832	phenol 2-monooxyge
32	0	0.0	4	2 A48360	gamma subunit of p
33	0	0.0	4	2 I40697	biotin A - Citroba
34	0	0.0	4	2 A61300	22K superhelical D
35	0	0.0	4	2 I57745	D-mannosate hydrol
36	0	0.0	4	2 A41890	protein D - Escher
37	0	0.0	4	2 S43014	hypothetical prote
38	0	0.0	4	2 D41654	hypothetical prote
39	0	0.0	4	2 B43848	cell surface adhes
40	0	0.0	4	2 I40505	hypothetical prote
41	0	0.0	4	2 I40870	phospholipase C (E
42	0	0.0	4	2 I40804	endoglucanase F -
43	0	0.0	4	2 T46627	hypothetical prote
44	0	0.0	4	2 S53508	starvation-induced
45	0	0.0	4	2 S17255	ribosomal protein

ALIGNMENTS

RESULT 1

RHTDIO  
thyroliberin - Bombina orientalis  
C;Species: Bombina orientalis  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A90919; A01415  
R;Yasuhara, T.; Nakajima, T.  
Chem. Pharm. Bull. 23, 3301-3303, 1975  
A;Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.  
A;Reference number: A90919; MUID:76138399; PMID:815011  
A;Accession: A90919  
A;Molecule type: protein  
A;Residues: 1-3 <YAS>  
C;Superfamily: thyroliberin precursor  
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 2

RHPGT  
thyroliberin - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A01415  
R;Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.  
Biochemistry 9, 1103-1106, 1970  
A;Title: Structure of porcine thyrotropin releasing hormone.  
A;Reference number: A90560; MUID:70136150; PMID:4984938  
A;Accession: A01415  
A;Molecule type: protein  
A;Residues: 1-3 <NAL>  
R;Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 37, 705-710, 1969  
A;Title: The identity of chemical and hormonal properties of the thyrotropin releasing hc  
A;Reference number: A90167; MUID:70039904; PMID:4982117  
A;Contents: annotation  
A;Note: biological activities and Rf values (in 17 chromatographic systems) of the synthe  
C;Superfamily: thyroliberin precursor  
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 3  
RHSH  
thyroliberin - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A93750; A01415  
R;Desiderio J.K., D.M.; Burgus, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.  
Org. Mass Spectrom. 5, 221-228, 1971  
A;Title: The elucidation of the primary structure of the hypothalamic thyroïd stimulating factor.  
A;Reference number: A93750  
A;Accession: A93750  
A;Molecule type: protein  
A;Residues: 1-3 <DES>  
R;Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.  
Nature 226, 321-325, 1970  
A;Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.  
A;Reference number: A93161; MUID:70163386; PMID:4985794  
A;Contents: annotation  
A;Note: physicochemical characteristics and biological activities of the natural and synthetic thyroliberin precursor  
C;Superfamily: thyroliberin precursor  
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 4  
A92971  
thyroliberin - eastern newt (tentative sequence)  
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A92971; A01415  
R;Grimm-Jorgensen, Y.; McKelvy, J.F.  
J. Neurochem. 23, 471-478, 1974  
A;Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain.  
A;Reference number: A92971; MUID:75035605; PMID:4214528  
A;Accession: A92971  
A;Molecule type: protein  
A;Residues: 1-3 <GRI>  
A;Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid  
C;Superfamily: thyroliberin precursor  
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 5  
GKHU  
growth-modulating peptide - human

C;Species: Homo sapiens (man)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A01421  
R;Schlesinger, D.H.; Pickart, L.; Thaler, M.M.  
Experientia 33, 324-325, 1977  
A;Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.  
A;Reference number: A01421; MUID:77162369; PMID:858356  
A;Accession: A01421  
A;Molecule type: protein  
A;Residues: 1-3 <SCH>  
A;Note: this serum tripeptide is found to stimulate growth of some cell types and to inhibit C;Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 G 1

RESULT 6  
A60898  
bursin - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A60898  
R;Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.  
Science 231, 997-999, 1986  
A;Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of the B-1 cell.  
A;Reference number: A60898; MUID:86122916; PMID:3484838  
A;Accession: A60898  
A;Molecule type: protein  
A;Residues: 1-3 <AUD>  
C;Superfamily: unassigned animal peptides  
C;Keywords: amidated carboxyl end; hormone  
F;3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 K 1

RESULT 7  
A23751  
spinal cord peptide SCP-4 - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C;Accession: A23751  
R;Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A;Reference number: A23751; MUID:85250425; PMID:4015098  
A;Accession: A23751  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-3 <HSI>  
C;Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 8

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

10M protein - protein search, using sw model

Run on: November 21, 2003, 15:20:03 ; Search time 2.97945 Seconds  
(without alignments)  
473.510 Million cell updates/sec

Title: US-09-010-377-4

Perfect score: 30

Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	1 GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	1 LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3	1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4	1 ACHI_ACHF	P35904 achatina fu
5	0	0.0	4	1 DCML_PSECH	P19916 pseudomonas
6	0	0.0	4	1 DCMS_PSECH	P19918 pseudomonas
7	0	0.0	4	1 EOSI_HUMAN	P02731 homo sapien
8	0	0.0	4	1 FAR3_HIRME	P42562 hirudo medi
9	0	0.0	4	1 FAR4_HIRME	P42563 hirudo medi
10	0	0.0	4	1 FFKA_ANTEL	P58705 anthopleura
11	0	0.0	4	1 FLRF_HIRME	P42561 hirudo medi
12	0	0.0	4	1 FLRN_ANTEL	P58707 anthopleura
13	0	0.0	4	1 FMRF_MACNI	P01162 macrocallis
14	0	0.0	4	1 FYRI_ANTEL	P58706 anthopleura
15	0	0.0	4	1 OCPI_OCTMI	P58648 octopus min
16	0	0.0	4	1 OCPS_OCTMI	P58649 octopus min
17	0	0.0	4	1 RM01_YEAST	P36515 saccharomyc
18	0	0.0	4	1 TUFT_HUMAN	P01858 homo sapien
19	0	0.0	5	1 ALI4_CARMA	P81817 carcinus ma
20	0	0.0	5	1 BIC0A_CITFR	P13071 citrobacter
21	0	0.0	5	1 BIC0B_CITFR	P12997 citrobacter
22	0	0.0	5	1 BPPT_BOTIN	P30425 bothrops in
23	0	0.0	5	1 EI03_LITRU	P82099 litoria rub
24	0	0.0	5	1 EI04_LITRU	P82100 litoria rub
25	0	0.0	5	1 FARP_ARTTR	P41853 artiopesthi
26	0	0.0	5	1 PAPT_PARMA	P81864 pardachirus
27	0	0.0	5	1 PRCT_PPRAM	P01373 periplaneta
28	0	0.0	5	1 PSK_DAUCA	P58261 daucus caro
29	0	0.0	5	1 RE11_LITRU	P82070 litoria rub
30	0	0.0	5	1 RE21_LITRU	P82071 litoria rub
31	0	0.0	5	1 RE31_LITRU	P82072 litoria rub
32	0	0.0	5	1 RE32_LITRU	P82073 litoria rub
33	0	0.0	5	1 SUGA_ACHDO	P19991 acheta dome

34	0	0.0	5	1 TPIS_CANFA	P54714 canis famil
35	0	0.0	5	1 TRM3_ECOLI	P13973 escherichia
36	0	0.0	5	1 UC22_MAIZE	P80628 zea mays (m
37	0	0.0	5	1 UF01_MOUSE	P38639 mus musculu
38	0	0.0	5	1 UXA4_CHLTR	P38005 chlamydia t
39	0	0.0	6	1 ACPH_RABIT	P25154 oryctolagus
40	0	0.0	6	1 ASP2_LACSN	P82655 lactobacill
41	0	0.0	6	1 CIP1_MYTED	P13736 mytilus edu
42	0	0.0	6	1 CIP2_MYTED	P13737 mytilus edu
43	0	0.0	6	1 EI01_LITRU	P82096 litoria rub
44	0	0.0	6	1 FARP_MONEX	P41966 moniezia ex
45	0	0.0	6	1 LOK1_LOCMI	P41491 locusta mig

## ALIGNMENTS

### RESULT 1

ID GRWM\_HUMAN STANDARD; PRT; 3 AA.  
AC P01157;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last annotation update)  
DE Growth-modulating peptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_SEQUENCE.  
RX MEDLINE=77162369; PubMed=858356;  
RA Schlesinger D.H., Pickart L., Thaler M.M.;  
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";  
RL Experimentia 33:324-325(1977).  
CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE  
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.  
DR GO; GO:0001558; P:regulation of cell growth; NAS.  
SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 G 1

### RESULT 2

ID LUXE\_VIBFI STANDARD; PRT; 3 AA.  
AC P24272;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Long-chain-fatty-acid-luciferin-component ligase (EC 6.2.1.19) (Acyl-  
protein synthetase) (Fragment).  
GN LUXE.  
OS Vibrio fischeri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
RN [1]\_SEQUENCE FROM N.A.  
RX MEDLINE=91072226; PubMed=2254256;  
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;  
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination  
site for the lux operon."  
RL J. Bacteriol. 172:6797-6802(1990).  
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE  
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS

```

CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
CC an acyl-protein thioester.
CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62812; -; NOT ANNOTATED_CDS.
CC DR LUMINESCENCE; Ligase.
CC KW NON TER 1 1
CC FT SEQUENCE 3 AA; 374 MW; 6AA3303000000000 CRC64;
CC SQ

```

Query Match 0.0%; Score 0; DB 1; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 1 1

```

RESULT 3
THYL PIG STANDARD; PRT; 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
OS Sus scrofa (Pig),
OS Ovis aries (Sheep),
OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
[1]_TaxID=9823, 9940, 8346, 8316;
SEQUENCE.
RC SPECIES=Pig; TISSUE=Hypothalamus;
RC MEDLINE=70136150; PubMed=4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RP SYNTHESIS.
RP SPECIES=Pig;
RC MEDLINE=70039904; PubMed=4982117;
RC Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
RA "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutamyl-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RP SPECIES=Sheep; TISSUE=Hypothalamus;
RC Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid
RT stimulating hormone releasing factor of ovine origin by means of mass
RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228(1971).
RN [4]
RP SYNTHESIS.
RC SPECIES=Sheep;
RC MEDLINE=70163386; PubMed=4985794;
RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA Guillemin R.;
RT "Characterization of ovine hypothalamic hypophysiotropic

```

```

RT TSH-releasing factor.";
RL Nature 226:321-325(1970).
RN [5]
RP SEQUENCE.
RC SPECIES=B.orientalis; TISSUE=Skin;
RX MEDLINE=76138399; PubMed=815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL Chem. Pharm. Bull. 23:3301-3303(1975).
RN [6]
RP SEQUENCE.
RC SPECIES=N.viridescens;
RX MEDLINE=75035605; PubMed=4214528;
RA Grimm-Joergensen Y., McKelvy J.F.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478(1974).
CC -!- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
CC PIR; A90919; RHTDIO.
DR PIR; A92971; A92971.
DR PIR; A93750; RSHST.
KW Amidation; Pyrrolidone carboxylic acid.
FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD RES 3 3 AMIDATION.
SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;

```

Query Match 0.0%; Score 0; DB 1; Length 3;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 Q 1

```

RESULT 4
ACH1 ACHFU STANDARD; PRT; 4 AA.
AC P35904;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Achatin-I.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC STRAIN=Perussac; TISSUE=Ganglion;
RX MEDLINE=89273551; PubMed=2597281;
RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
RT fulica Perussac containing a D-amino acid residue.";
RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
RN [2]
RP CHARACTERIZATION.
RC STRAIN=Perussac; TISSUE=Heart atrium;
RX MEDLINE=91264856; PubMed=1675568;
RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
RT "Purification of achatin-I from the atria of the African giant snail,
RT Achatina fulica, and its possible function.";
RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=93014529; PubMed=1399265;
RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 15:20:39 : Search time 12.9452 Seconds  
(without alignments)  
598.027 Million cell updates/sec

Title: US-09-010-377-4  
Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXX 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*  
1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	% Query Match	Length	ID	Description
1	0	0.0	4	11 Q08433	Q08433 rattus sp.
2	0	0.0	5	2 P83073	P83073 bacillus ce
3	0	0.0	5	10 Q99007	Q99007 hordeum vul
4	0	0.0	5	13 P83308	P83308 gallus gall
5	0	0.0	6	10 P82181	P82181 spinacia ol
6	0	0.0	6	10 P82541	P82541 spinacia ol
7	0	0.0	6	10 P82182	P82182 spinacia ol
8	0	0.0	7	2 Q8KMS3	Q8KMS3 klebsiella
9	0	0.0	7	2 Q47505	Q47505 escherichia
10	0	0.0	7	2 P70804	P70804 azotobacter
11	0	0.0	7	2 O50556	O50556 actinobacil
12	0	0.0	7	2 Q47477	Q47477 escherichia
13	0	0.0	7	2 Q47029	Q47029 enterobacte
14	0	0.0	7	2 P72081	P72081 nocardia la
15	0	0.0	7	2 Q54248	Q54248 streptomyce
16	0	0.0	7	2 Q8KMS9	Q8KMS9 enterobacte

17	0	0.0	7	2	O34028	sphingomona
18	0	0.0	7	2	O07354	synechococc
19	0	0.0	7	2	Q8GL12	borrelia bu
20	0	0.0	7	2	Q8GL04	borrelia bu
21	0	0.0	7	2	Q8GL00	borrelia bu
22	0	0.0	7	4	Q15903	Q15903 homo sapien
23	0	0.0	7	4	Q8NH7	Q8NH7 homo sapien
24	0	0.0	7	4	Q15897	Q15897 homo sapien
25	0	0.0	7	6	Q28742	Q28742 oryctolagus
26	0	0.0	7	8	P92214	P92214 amblyopium
27	0	0.0	7	8	P92393	P92393 hordeum vul
28	0	0.0	7	8	P92403	P92403 lophopyrum
29	0	0.0	7	8	P92427	P92427 peridictyon
30	0	0.0	7	8	O99182	O99182 gnatholebia
31	0	0.0	7	8	P92430	P92430 aegilops ta
32	0	0.0	7	8	P92221	P92221 bromus iner
33	0	0.0	7	8	O98866	O98866 spinacia ol
34	0	0.0	7	8	P92425	P92425 pseudoroegn
35	0	0.0	7	8	P92381	P92381 hordeum bra
36	0	0.0	7	8	P92387	P92387 henrardia p
37	0	0.0	7	8	P92210	P92210 agropyron c
38	0	0.0	7	8	P92440	P92440 thinopyrum
39	0	0.0	7	8	P92218	P92218 australopyr
40	0	0.0	7	8	P92390	P92390 heteranthe
41	0	0.0	7	8	P92372	P92372 haynaldia v
42	0	0.0	7	8	P92442	P92442 taeniathe
43	0	0.0	7	8	P92226	P92226 crithopsis
44	0	0.0	7	8	Q8MFY6	Q8MFY6 taraxacum (
45	0	0.0	7	8	Q95945	Q95945 saccharomyc

#### ALIGNMENTS

#### RESULT 1

Q08433  
ID Q08433 PRELIMINARY; PRT; 4 AA.  
AC Q08433;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Bilirubin UDP-glucuronosyltransferase (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_taxid=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Gunn;  
RX MEDLINE=91282758; PubMed=1840486;  
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;  
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the  
hyperbilirubinemic Gunn rat."  
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
DR EMBL; S38636; AAB19259.1; -.  
KW Transferase.  
FT NON\_TER 1 1  
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 N 1

#### RESULT 2

P83073  
ID P83073 PRELIMINARY; PRT; 5 AA.  
AC P83073;  
DT 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE 88 kDa protein (Fragment).  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 NCBI\_TaxID=1396;  
 (1)  
 RN RRNA  
 RP SEQUENCE  
 RC STRAIN=NCIMB 11796;  
 RA Browne N., Dows B.C.A.;  
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
 FT NON TER 5  
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;  
 Query Match 0.0%; Score 0; DB 2; Length 5;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 4 T 4

RESULT 3  
 Q99007  
 ID Q99007 PRELIMINARY; PRT; 5 AA.  
 AC Q99007;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Alpha amylase (Fragment).  
 GN AMY1 GENE.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Hordeum.  
 NCBI\_TaxID=4513;  
 (1)  
 RN RRNA  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91329704; PubMed=1831055;  
 RA Jacobsen J.V., Close T.J.;  
 RT "Control of transient expression of chimaeric genes by gibberellic acid and abscisic acid in protoplasts prepared from mature barley aleurone layers."  
 RT Plant Mol. Biol. 16:713-721(1991).  
 DR ENBL; X54643; CAA38455.1; -.  
 FT NON TER 5  
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
 Query Match 0.0%; Score 0; DB 10; Length 5;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 2 A 2

RESULT 4  
 P83308  
 ID P83308 PRELIMINARY; PRT; 5 AA.  
 AC P83308;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE FMRFamide-like neuropeptide (1PLRF-amide).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 NCBI\_TaxID=9031;  
 (1)  
 RN RRNA  
 RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Brain;  
 RX PubMed=6137771;  
 RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;  
 RT "A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide";  
 RL Nature 305:328-330(1983).  
 CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.  
 KW Neuropeptide.  
 SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;  
 Query Match 0.0%; Score 0; DB 13; Length 5;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 L 1

RESULT 5  
 P82181  
 ID P82181 PRELIMINARY; PRT; 6 AA.  
 AC P82181;  
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 NCBI\_TaxID=3562;  
 (1)  
 RN RRNA  
 RP SEQUENCE.  
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro; IPR001790; Ribosomal L10.  
 DR InterPro; IPR002363; Ribosomal\_L10sub.  
 DR Pfam; PF00466; Ribosomal\_L10; PARTIAL.  
 DR PROSITE; PS01109; RIBOSOMAL\_L10; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON TER 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB0000 CRC64;  
 Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 A 1

RESULT 6  
 P82541  
 ID P82541 PRELIMINARY; PRT; 6 AA.  
 AC P82541;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
 OS Spinacia oleracea (Spinach).



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

MM protein - protein search, using sw model

Run on: November 21, 2003, 15:19:18 ; Search time 9.9863 Seconds  
(without alignments)  
286.100 Million cell updates/sec

Hit: 35  
Perfect score: 35  
Sequence: 1 XXXXXXXPVFXXXXXX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_15Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	60.0	4	17 AAR92613	VLA-4 binding pept
2	21	60.0	5	17 AAR92607	VLA-4 binding pept
3	21	60.0	6	17 AAR92604	VLA-4 binding pept
4	21	60.0	6	17 AAR92614	VLA-4 binding pept
5	21	60.0	6	17 AAR92531	VLA-4 binding pept
6	21	60.0	7	17 AAR92599	VLA-4 binding pept
7	21	60.0	8	17 AAR92602	VLA-4 binding pept
8	21	60.0	9	17 AAR92610	VLA-4 binding pept
9	21	60.0	9	22 ABP21019	HIV A03 motif pol

10	21	60.0	10	17 AAR92598	VLA-4 binding pept
11	21	60.0	10	17 AAR92573	VLA-4 binding pept
12	21	60.0	10	21 AAY56008	Mucor hiemalis end
13	21	60.0	10	22 ABP21021	HIV A03 motif pol
14	21	60.0	11	17 AAR92563	VLA-4 binding pept
15	21	60.0	11	17 AAR92601	VLA-4 binding pept
16	21	60.0	11	22 ABP21023	HIV A03 motif pol
17	21	60.0	12	21 AAY93861	Reactive peptide w
18	21	60.0	13	22 AAB49917	Human P-glycoprote
19	21	60.0	13	22 AAB49920	Human P-glycoprote
20	21	60.0	14	23 ABG93568	Human P-glycoprote
21	21	60.0	14	23 ABG93704	Human P-glycoprote
22	21	60.0	14	23 ABP47053	Human Blys binding
23	21	60.0	15	23 ABP58590	Human cysteine swi
24	21	60.0	15	23 AAO14991	Human synaptophys
25	21	60.0	18	21 AAY91149	Modified HBV surfa
26	21	60.0	18	23 ABG68218	Hepatitis B virus
27	21	60.0	19	21 AAB34287	Human secreted pro
28	21	60.0	22	21 AAY65274	Human 5' EST relat
29	21	60.0	25	24 ABJ18878	Human Bcl-XL-bindi
30	21	60.0	27	22 AAM95657	Human reproductive
31	21	60.0	28	21 AAB45151	Human secreted pro
32	21	60.0	30	21 AAY91189	Modified HBV surfa
33	21	60.0	32	22 ABG48897	Human liver peptid
34	21	60.0	32	22 ABG53276	Human liver peptid
35	21	60.0	32	22 ABP28892	Peptide #1543 enco
36	21	60.0	32	22 ABP34068	Peptide #1574 enco
37	21	60.0	32	22 ABP38450	Peptide #5956 enco
38	21	60.0	32	22 ABP19508	Protein #1507 enco
39	21	60.0	32	22 ABP23617	Protein #5616 enco
40	21	60.0	32	22 AAM54850	Human brain expres
41	21	60.0	32	22 AAM59061	Human brain expres
42	21	60.0	32	22 AAM67232	Human bone marrow
43	21	60.0	32	22 AAM71594	Human bone marrow
44	21	60.0	32	22 AAM15077	Peptide #1511 enco
45	21	60.0	32	22 AAM27525	Peptide #1562 enco

ALIGNMENTS

RESULT 1				
AAR92613				
ID	AAR92613	standard; peptide; 4 AA.		
XX				
AC	AAR92613;			
XX				
DT	10-SEP-1996	(first entry)		
XX				
DE	VLA-4 binding peptide #73.			
XX				
KW	VLA-4; inhibitor; binding peptide; leucocyte adhesion; CD49d/CD29; CDR; alpha4beta1 integrin; VCAM-1; endothelial cell; antibody; meningitis;			
KW	complementary determining region; inflammatory brain disorder; therapy;			
KW	multiple sclerosis; encephalitis; Alzheimer's disease; atherosclerosis;			
KW	AIDS dementia; diabetes; inflammatory bowel disease; tumour metastases;			
KW	rheumatoid arthritis; transplant rejection; myocardial ischaemia; asthma.			
XX	Synthetic.			
OS				
XX				
FH	Key	Location/Qualifiers		
FT	Modified-site	1	/note= "acetylated"	
XX				
PN	WO9601644-A1.			
XX				
PD	25-JAN-1996.			
XX				
PF	10-JUL-1995;	95WO-US08516.		
XX				
PR	06-JUN-1995;	95US-0467580.		
PR	11-JUL-1994;	94US-0273055.		
XX				

```

PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Pleiss MA, Thorsett ED, Yednock TA;
PI WPI; 1996-097452/10.
XX
XX New peptide(s) that bind VLA-4, inhibit leucocyte adhesion - useful
PT esp. for treatment of inflammatory disease, e.g. multiple sclerosis
XX
XX Example 9; Page 29; 42pp; English.
PS
XX AAR92598-R92659 represents VLA-4 binding peptides derived from the
XX peptides shown in AAR92536-R92555 and AAR92563-R92594. VLA-4 is also
CC known as alpha4beta1 integrin and CD49d/CD29. VLA-4 binds to the
CC non-matrix molecule VCAM-1, which is expressed by endothelial and other
CC cells. These sequences mimic short consensus sequences in the
CC complementarity determining region 3 of antibody 21/6, and the FGN region
CC of VCAM-1. These sequences inhibit the adhesion of leucocytes mediated by
CC VLA-4. It is useful for treating inflammatory brain disorders
CC (especially multiple sclerosis), meningitis, encephalitis, asthma,
CC Alzheimer's disease, atherosclerosis, AIDS dementia, diabetes,
CC inflammatory bowel disease, rheumatoid arthritis, transplant rejection,
CC tumour metastases and myocardial ischaemia. These sequences can also be
CC labelled, and can then be used for in vivo or in vitro diagnosis, such as
CC monitoring inflammatory responses, isolating leucocytes, in assays for
CC inhibitors of VLA-4/VCAM-1 interactions and imaging sites of
CC inflammation.
XX
XX Sequence 4 AA;
SQ Query Match 60.0%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 1 PVSF 4

RESULT 2
AAR92607
ID AAR92607 standard; peptide; 5 AA.
XX
XX AAR92607;
XX
XX 10-SEP-1996 (first entry)
DT
DE VLA-4 binding peptide #67.
XX
XX VLA-4; inhibitor; binding peptide; leucocyte adhesion; CD49d/CD29; CDR;
KW alpha4beta1 integrin; VCAM-1; endothelial cell; antibody; meningitis;
KW complementarity determining region; inflammatory brain disorder; therapy;
KW multiple sclerosis; encephalitis; Alzheimer's disease; atherosclerosis;
KW AIDS dementia; diabetes; inflammatory bowel disease; tumour metastases;
KW rheumatoid arthritis; transplant rejection; myocardial ischaemia; asthma.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "acetylated"
FT
XX
XX WO9601644-A1.
XX
XX 25-JAN-1996.
XX
XX 10-JUL-1995; 95WO-US08516.
XX
XX 06-JUN-1995; 95US-0467580.
XX
XX 11-JUL-1994; 94US-0273055.
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Pleiss MA, Thorsett ED, Yednock TA;
XX

```

```

PI Pleiss MA, Thorsett ED, Yednock TA;
XX
XX WPI; 1996-097452/10.
XX
XX New peptide(s) that bind VLA-4, inhibit leucocyte adhesion - useful
PT esp. for treatment of inflammatory disease, e.g. multiple sclerosis
XX
XX Example 9; Page 29; 42pp; English.
PS
XX AAR92598-R92659 represents VLA-4 binding peptides derived from the
XX peptides shown in AAR92536-R92555 and AAR92563-R92594. VLA-4 is also
CC known as alpha4beta1 integrin and CD49d/CD29. VLA-4 binds to the
CC non-matrix molecule VCAM-1, which is expressed by endothelial and other
CC cells. These sequences mimic short consensus sequences in the
CC complementarity determining region 3 of antibody 21/6, and the FGN region
CC of VCAM-1. These sequences inhibit the adhesion of leucocytes mediated by
CC VLA-4. It is useful for treating inflammatory brain disorders
CC (especially multiple sclerosis), meningitis, encephalitis, asthma,
CC Alzheimer's disease, atherosclerosis, AIDS dementia, diabetes,
CC inflammatory bowel disease, rheumatoid arthritis, transplant rejection,
CC tumour metastases and myocardial ischaemia. These sequences can also be
CC labelled, and can then be used for in vivo or in vitro diagnosis, such as
CC monitoring inflammatory responses, isolating leucocytes, in assays for
CC inhibitors of VLA-4/VCAM-1 interactions and imaging sites of
CC inflammation.
XX
XX Sequence 5 AA;
SQ Query Match 60.0%; Score 21; DB 17; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 2 PVSF 5

RESULT 3
AAR92604
ID AAR92604 standard; peptide; 6 AA.
XX
XX AAR92604;
XX
XX 10-SEP-1996 (first entry)
DT
DE VLA-4 binding peptide #64.
XX
XX VLA-4; inhibitor; binding peptide; leucocyte adhesion; CD49d/CD29; CDR;
KW alpha4beta1 integrin; VCAM-1; endothelial cell; antibody; meningitis;
KW complementarity determining region; inflammatory brain disorder; therapy;
KW multiple sclerosis; encephalitis; Alzheimer's disease; atherosclerosis;
KW AIDS dementia; diabetes; inflammatory bowel disease; tumour metastases;
KW rheumatoid arthritis; transplant rejection; myocardial ischaemia; asthma.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "acetylated"
FT
XX
XX WO9601644-A1.
XX
XX 25-JAN-1996.
XX
XX 10-JUL-1995; 95WO-US08516.
XX
XX 06-JUN-1995; 95US-0467580.
XX
XX 11-JUL-1994; 94US-0273055.
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Pleiss MA, Thorsett ED, Yednock TA;
XX

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: November 21, 2003, 15:21:54 ; Search time 3.39041 Seconds  
(without alignments)  
224.632 Million cell updates/sec

title: US-09-010-377-5

effect score: 35

sequence: 1 XXXXXXXPVSFXXXXXX 18

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	60.0	4	3	US-08-467-580-82
2	21	60.0	4	5	PCT-US95-08516-82
3	21	60.0	5	3	US-08-467-580-75
4	21	60.0	5	5	PCT-US95-08516-75
5	21	60.0	6	3	US-08-467-580-72
6	21	60.0	6	3	US-08-467-580-83
7	21	60.0	6	5	PCT-US95-08516-72
8	21	60.0	6	5	PCT-US95-08516-83
9	21	60.0	7	3	US-08-467-580-67
10	21	60.0	7	5	PCT-US95-08516-67
11	21	60.0	8	3	US-08-467-580-70
12	21	60.0	8	5	PCT-US95-08516-70
13	21	60.0	9	3	US-08-467-580-79
14	21	60.0	9	5	PCT-US95-08516-79
15	21	60.0	10	3	US-08-467-580-40
16	21	60.0	10	3	US-08-467-580-66
17	21	60.0	10	3	US-08-467-580-87
18	21	60.0	10	5	PCT-US95-08516-40
19	21	60.0	10	5	PCT-US95-08516-66
20	21	60.0	10	5	PCT-US95-08516-87
21	21	60.0	11	3	US-08-467-580-30
22	21	60.0	11	3	US-08-467-580-69
23	21	60.0	11	5	PCT-US95-08516-30
24	21	60.0	11	5	PCT-US95-08516-69
25	21	60.0	18	3	US-09-100-414B-29
26	21	60.0	18	3	US-09-303-323-29
27	21	60.0	18	4	US-09-770-014-29

28	21	60.0	30	3	US-09-100-414B-69	Sequence 69, Appl
29	21	60.0	30	3	US-09-303-323-69	Sequence 69, Appl
30	21	60.0	30	4	US-09-770-014-69	Sequence 69, Appl
31	21	60.0	43	4	US-09-369-247-122	Sequence 122, Appl
32	21	60.0	45	3	US-08-975-080-11	Sequence 11, Appl
33	21	60.0	46	1	US-08-340-428B-36	Sequence 36, Appl
34	21	60.0	46	4	US-09-205-258-1058	Sequence 1058, Appl
35	21	60.0	46	5	PCT-US93-07306-36	Sequence 36, Appl
36	21	60.0	48	4	US-09-904-615-74	Sequence 74, Appl
37	21	60.0	67	2	US-08-511-485-14	Sequence 14, Appl
38	21	60.0	67	4	US-09-201-936-14	Sequence 14, Appl
39	21	60.0	86	4	US-09-134-001C-5329	Sequence 5329, Appl
40	21	60.0	92	1	US-08-261-304-4	Sequence 4, Appl
41	21	60.0	92	1	US-08-261-304-5	Sequence 5, Appl
42	21	60.0	92	2	US-08-447-464-6	Sequence 6, Appl
43	21	60.0	92	2	US-08-716-679-6	Sequence 6, Appl
44	21	60.0	94	4	US-09-252-991A-21994	Sequence 21994, A
45	21	60.0	96	1	US-08-313-681A-5	Sequence 5, Appl

## ALIGNMENTS

### RESULT 1

US-08-467-580-82

; Sequence 82, Application US/08467580B

; Patent No. 6001809

; GENERAL INFORMATION:

; APPLICANT: Thorsett, Eugene D

; APPLICANT: Vednock, Theodore A

; APPLICANT: Pleiss, Michael A

; TITLE OF INVENTION: Inhibitors of Leukocyte Adhesion

; FILE REFERENCE: 123-US-CIP1

; CURRENT APPLICATION NUMBER: US/08/467,580B

; CURRENT FILING DATE: 1995-06-06

; EARLIER APPLICATION NUMBER: 08/273,055

; EARLIER FILING DATE: 1994-07-11

; NUMBER OF SEQ ID NOS: 163

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 82

; LENGTH: 4

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: MOD\_RES

; LOCATION: {1}

; OTHER INFORMATION: ACETYLATION

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Peptide

; OTHER INFORMATION: derived from VCAM-1

US-08-467-580-82

Query Match 60.0%; Score 21; DB 3; Length 4;

Best Local Similarity 100.0%; Pred.No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PVSF 11

Db 1 PVSF 4

### RESULT 2

PCT-US95-08516-82

; Sequence 82, Application PC/TUS9508516

; GENERAL INFORMATION:

; APPLICANT: ATHENA NEUROSCIENCES, INC.

; TITLE OF INVENTION: INHIBITORS OF LEUKOCYTE ADHESION

; NUMBER OF SEQUENCES: 157

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ATHENA NEUROSCIENCES, INC.

; STREET: 800 Gateway Blvd.

; CITY: South San Francisco

; STATE: CA

COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08516  
FILING DATE: 10-JUL-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/273,055  
FILING DATE: 11-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: MOOI, LESLIE A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 002010-008  
TELEPHONE: (415) 877-0900  
TELEFAX: (415) 877-3620  
INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US95-08516-82

Query Match 60.0%; Score 21; DB 5; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSP 11  
|||||

Db 1 PVSP 4

RESULT 3  
US-08-467-580-75  
; Sequence 75, Application US/08467580B  
; Patent No. 6001809  
; GENERAL INFORMATION:  
; APPLICANT: Thorsett, Eugene D  
; APPLICANT: Vednock, Theodore A  
; APPLICANT: Pleiss, Michael A  
; TITLE OF INVENTION: Inhibitors of Leukocyte Adhesion  
; FILE REFERENCE: 123-US-CIPI  
; CURRENT APPLICATION NUMBER: US/08/467,580B  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/273,055  
; EARLIER FILING DATE: 1994-07-11  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 75  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: ACETYLATION  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide  
; OTHER INFORMATION: derived from VCAM-1  
US-08-467-580-75

Query Match 60.0%; Score 21; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSP 11  
|||||

Db 2 PVSP 5

RESULT 4  
PCT-US95-08516-75  
; Sequence 75, Application PC/TUS9508516  
; GENERAL INFORMATION:  
; APPLICANT: ATHENA NEUROSCIENCES, INC.  
; TITLE OF INVENTION: INHIBITORS OF LEUKOCYTE ADHESION  
; NUMBER OF SEQUENCES: 157  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ATHENA NEUROSCIENCES, INC.  
; STREET: 800 Gateway Blvd.  
; CITY: South San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08516  
; FILING DATE: 10-JUL-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/273,055  
; FILING DATE: 11-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MOOI, LESLIE A.  
; REGISTRATION NUMBER: 37,047  
; REFERENCE/DOCKET NUMBER: 002010-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 877-0900  
; TELEFAX: (415) 877-3620  
; INFORMATION FOR SEQ ID NO: 75:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
PCT-US95-08516-75

Query Match 60.0%; Score 21; DB 5; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSP 11  
|||||

Db 2 PVSP 5

RESULT 5  
US-08-467-580-72  
; Sequence 72, Application US/08467580B  
; Patent No. 6001809  
; GENERAL INFORMATION:  
; APPLICANT: Thorsett, Eugene D  
; APPLICANT: Vednock, Theodore A  
; APPLICANT: Pleiss, Michael A  
; TITLE OF INVENTION: Inhibitors of Leukocyte Adhesion  
; FILE REFERENCE: 123-US-CIPI  
; CURRENT APPLICATION NUMBER: US/08/467,580B  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/273,055  
; EARLIER FILING DATE: 1994-07-11  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 72  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: November 21, 2003, 15:27:54 ; Search time 6.71918 Seconds  
(without alignments)  
489.058 Million cell updates/sec

Title: US-09-010-377-5

Perfect score: 35

Sequence: 1 XXXXXPFVFXXXXXX 18

Scoring table:

Blosum62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	60.0	11	US-10-062-710-100	Sequence 100, App
2	21	60.0	12	US-09-990-832C-103	Sequence 103, App
3	21	60.0	13	US-10-010-310-3	Sequence 3, Appli
4	21	60.0	13	US-10-010-310-6	Sequence 6, Appli
5	21	60.0	14	US-09-982-172-88	Sequence 88, Appl
6	21	60.0	14	US-09-982-172-224	Sequence 224, App
7	21	60.0	14	US-09-880-748-3064	Sequence 3064, App
8	21	60.0	18	US-09-747-802-65	Sequence 65, Appl
9	21	60.0	18	US-09-865-294-58	Sequence 58, Appl
10	21	60.0	25	US-10-082-750-57	Sequence 57, Appl
11	21	60.0	27	US-09-764-891-4315	Sequence 4315, Ap
12	21	60.0	32	US-09-864-761-34806	Sequence 34806, A
13	21	60.0	32	US-09-864-761-38915	Sequence 38915, A
14	21	60.0	33	US-10-144-929-92	Sequence 92, Appl
15	21	60.0	35	US-09-764-877-1329	Sequence 1329, Ap

16	21	60.0	41	9	US-09-864-761-34763	Sequence 34763, A
17	21	60.0	42	9	US-09-864-761-42139	Sequence 42139, A
18	21	60.0	43	9	US-09-864-761-34242	Sequence 34242, A
19	21	60.0	43	12	US-10-029-386-33238	Sequence 33238, A
20	21	60.0	43	15	US-10-050-704-136	Sequence 136, App
21	21	60.0	43	15	US-10-062-548-122	Sequence 122, App
22	21	60.0	44	9	US-09-864-761-47048	Sequence 47048, A
23	21	60.0	44	9	US-09-864-761-47765	Sequence 47765, A
24	21	60.0	44	11	US-09-764-891-4075	Sequence 4075, Ap
25	21	60.0	45	9	US-09-864-761-44905	Sequence 44905, A
26	21	60.0	45	15	US-10-138-618-11	Sequence 11, Appl
27	21	60.0	46	12	US-09-933-767-1058	Sequence 1058, Ap
28	21	60.0	46	15	US-10-023-282-1058	Sequence 1058, Ap
29	21	60.0	47	11	US-09-764-891-2821	Sequence 2821, Ap
30	21	60.0	48	9	US-09-739-254-74	Sequence 74, Appl
31	21	60.0	48	9	US-09-904-615-74	Sequence 74, Appl
32	21	60.0	48	12	US-10-055-098-74	Sequence 74, Appl
33	21	60.0	48	12	US-10-029-386-31823	Sequence 31823, A
34	21	60.0	48	15	US-10-054-988-74	Sequence 74, Appl
35	21	60.0	49	9	US-09-764-878-112	Sequence 112, App
36	21	60.0	49	15	US-10-079-854-112	Sequence 112, App
37	21	60.0	50	9	US-09-864-761-45925	Sequence 45925, A
38	21	60.0	52	9	US-09-864-761-37265	Sequence 37265, A
39	21	60.0	52	9	US-09-864-761-47325	Sequence 47325, A
40	21	60.0	52	11	US-09-764-872-246	Sequence 246, App
41	21	60.0	53	9	US-09-864-761-46816	Sequence 46816, A
42	21	60.0	54	9	US-09-925-297-466	Sequence 466, App
43	21	60.0	54	12	US-10-012-952A-245	Sequence 245, App
44	21	60.0	60	11	US-09-764-891-5136	Sequence 5136, App
45	21	60.0	60	12	US-10-029-386-30343	Sequence 30343, A

ALIGNMENTS

RESULT 1

US-10-062-710-100

; Sequence 100, Application US/10062710

; Publication No. US20030049253A1

; GENERAL INFORMATION:

; APPLICANT: Li, Frank Q.

; APPLICANT: Chu, Yong-Liang

; APPLICANT: Qid, Jian-Tai

; TITLE OF INVENTION: Polymeric Conjugates for Delivery of

; TITLE OF INVENTION: MHC-Recognized Epitopes

; TITLE OF INVENTION: Via Peptide Vaccines

; FILE REFERENCE: 3781-001-27

; CURRENT APPLICATION NUMBER: US/10/062,710

; CURRENT FILING DATE: 2002-02-05

; PRIOR APPLICATION NUMBER: US 60/310,498

; PRIOR FILING DATE: 2001-08-08

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 100

; LENGTH: 11

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HIV CTL-Epitopes

US-10-062-710-100

Query Match 60.0%; Score 21; DB 15; Length 11;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PVSF 11

Db 1 PVSF 4

RESULT 2

US-09-990-832C-103

; Sequence 103, Application US/09990832C

```
, Publication No. US20030149235A1
; GENERAL INFORMATION:
; APPLICANT: University Court of the University of Glasgow
; TITLE OF INVENTION: Targeting peptides
; FILE REFERENCE: PC/MC/JM/Fl1910US
; CURRENT APPLICATION NUMBER: US/09/990,832C
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Targeting peptide sequence
US-09-990-832C-103

Query Match 60.0%; Score 21; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 6 PVSF 9

RESULT 3
US-10-010-310-3
; Sequence 3, Application US/10010310
; Publication No. US20020142348A1
; GENERAL INFORMATION:
; APPLICANT: MCGILL UNIVERSITY
; APPLICANT: GEORGES, Elias
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTIONS AND METHODS FOR
; TITLE OF INVENTION: IDENTIFYING INTERACTING PROTEINS AND THE AMINO ACID
; FILE REFERENCE: 112418.122
; CURRENT APPLICATION NUMBER: US/10/010,310
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/134,259
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Targeting peptide sequence
US-10-010-310-3

Query Match 60.0%; Score 21; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 1 PVSF 4

RESULT 4
US-10-010-310-6
; Sequence 6, Application US/10010310
; Publication No. US20020142348A1
; GENERAL INFORMATION:
; APPLICANT: MCGILL UNIVERSITY
; APPLICANT: GEORGES, Elias
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTIONS AND METHODS FOR
; TITLE OF INVENTION: IDENTIFYING INTERACTING PROTEINS AND THE AMINO ACID
; FILE REFERENCE: 112418.122
; CURRENT APPLICATION NUMBER: US/10/010,310
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/134,259
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Targeting peptide sequence
US-10-010-310-6

Query Match 60.0%; Score 21; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 1 PVSF 4

RESULT 5
US-09-982-172-88
; Sequence 88, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATI
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-88

Query Match 60.0%; Score 21; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 9 PVSF 12

RESULT 6
US-09-982-172-224
; Sequence 224, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATI
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 224
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-224

Query Match 60.0%; Score 21; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 5.1.6  
Copyright (C) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

run on: November 21, 2003, 15:20:03 ; Search time 1.78767 Seconds  
(without alignments)  
473.510 Million cell updates/sec

id: us-09-010-377-5  
effect score: 35  
sequence: 1 XXXXXXPFVFXXXXXX 18

coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	21	60.0	57	YOR3_TTV1	P19278 thermoprote
2	21	60.0	69	YAT1_RHOBL	P05443 rhodopseudo
3	21	60.0	84	RLX7_STRAA	Q53602 staphylococ
4	21	60.0	84	Y353_METJA	O06918 methanococ
5	21	60.0	88	YVAD_VACCC	P20513 vaccinia vi
6	21	60.0	96	ICTL_PIG	P15175 sus scrofa
7	21	60.0	99	YL30_ARCFU	Q28150 archaeglob
8	21	60.0	104	YBK0_YEAST	P38190 saccharomyc
9	21	60.0	107	LRP2_HSVIF	P17589 herpes simp
10	21	60.0	122	FOLB_CHLMU	Q9PJC6 chlamydia m
11	21	60.0	124	FOLB_CHLTR	O84620 chlamydia t
12	21	60.0	124	YBF3_YEAST	P38190 saccharomyc
13	21	60.0	125	OM10_RHILO	Q98h63 rhizobium l
14	21	60.0	139	Y410_RHIME	Q52911 rhizobium m
15	21	60.0	141	RL11_LIBAS	P36250 liberibacte
16	21	60.0	144	INDC_BOVIN	P33046 bos taurus
17	21	60.0	144	YLX3_CAEEL	P46499 caenorhabdi
18	21	60.0	147	PG2_PIG	P32195 sus scrofa
19	21	60.0	147	R157_BOVIN	Q28183 bos taurus
20	21	60.0	147	Y720_SYNY3	P74650 synectocyst
21	21	60.0	149	PG1_PIG	P32194 sus scrofa
22	21	60.0	149	PG3_PIG	P32196 sus scrofa
23	21	60.0	149	PG4_PIG	P49933 sus scrofa
24	21	60.0	149	PG5_PIG	P49934 sus scrofa
25	21	60.0	149	YG2U_YEAST	P53263 saccharomyc
26	21	60.0	153	CORZ_COWRU	Q04302 cowdria rum
27	21	60.0	153	MP23_PIG	P49930 sus scrofa
28	21	60.0	153	R1B1_CAUCR	Q9A9S4 caulobacter
29	21	60.0	154	R1B1_CAMJE	Q9P1B9 campylobact
30	21	60.0	155	Y654_CHLPN	Q9Z7Q0 chlamydia p
31	21	60.0	156	R1B8_HELPJ	Q9Zn56 helicobacte
32	21	60.0	156	R1B8_HELPJ	Q24854 helicobacte
33	21	60.0	158	MB27_BOVIN	P54228 bos taurus

34	21	60.0	159	1	MB28_BOVIN	P54229 bos taurus
35	21	60.0	160	1	SC53_SHEEP	P49928 ovis aries
36	21	60.0	160	1	SC52_SHEEP	P49929 ovis aries
37	21	60.0	163	1	TPX1_STRPN	P72500 streptococc
38	21	60.0	165	1	MB34_BOVIN	P56425 bos taurus
39	21	60.0	166	1	MP36_PIG	P49931 sus scrofa
40	21	60.0	167	1	MP37_PIG	P49932 sus scrofa
41	21	60.0	170	1	PA39_HUMAN	P49913 homo sapien
42	21	60.0	171	1	CP18_RABIT	P25230 oryctolagus
43	21	60.0	172	1	PR39_PIG	P80054 sus scrofa
44	21	60.0	174	1	FANG_ECOLI	P20861 escherichia
45	21	60.0	176	1	ECT5_BOVIN	P19660 bos taurus

ALIGNMENTS

RESULT 1

YOR3\_TTV1  
ID\_YOR3\_TTV1 STANDARD; PRT; 57 AA.  
AC P19278;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-NOV-1990 (Rel. 16, Last annotation update)  
DE Hypothetical 6.9 kDa protein.  
OS Thermoproteus tenax virus 1 (strain KRA1) (TTV1).  
OC Viruses; dsDNA viruses, no RNA stage; Lipothrrixviridae;  
OC Lipothrrixvirus.  
OX NCBI\_TaxID=10480;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Neumann H.;  
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; X14855; CAA32971.1; -  
DR Hypothetical protein.  
KW Hypothetical protein.  
SQ SEQUENCE 57 AA; 6914 MW; B907370704263763 CRC64;

Query Match 60.0%; Score 21; DB 1; Length 57;  
Best Local Similarity 100.0%; Pred.No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11  
Dd 6 PVSF 9

RESULT 2

YAT1\_RHOBL  
ID\_YAT1\_RHOBL STANDARD; PRT; 69 AA.  
AC P05443;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-NOV-1988 (Rel. 09, Last annotation update)  
DE ATP synthase subunits region ORF 1 (fragment).  
OS Rhodopseudomonas blattica.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;  
OC Rhodobacteraceae; Rhodobacter.  
OX NCBI\_TaxID=1075;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85058188; PubMed=6209404;  
RA Tybulewicz V.L.J., Falk G., Walker J.E.;  
RT "Rhodopseudomonas blattica atp operon. Nucleotide sequence and  
transcription.";

```

RL J. Mol. Biol. 179:185-214(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z00018; CAA77307.1; -.
DR DR
DR PIR; S04666; S04666.
DR KW Hypothetical protein.
FT NON TER 1
FT SEQUENCE 69 AA; 7471 MW; 574EB0B6F8529ED9 CRC64;
CC -----
CC Query Match 60.0%; Score 21; DB 1; Length 69;
CC Best Local Similarity 100.0%; Pred. No. 74;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 8 PVSF 11
CC Db 26 PVSF 29
CC -----
CC RESULT 3
CC RXL7 STAM STANDARD; PRT; 84 AA.
CC ID RXL7 STAM
CC AC Q53602; Q99W63;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Putative ribosomal protein L7Ae-like.
CC GN SAV0544 OR SA0502 OR MW0499.
CC OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
CC OS Staphylococcus aureus (strain N315),
CC OS Staphylococcus aureus (strain MW2), and
CC OS Staphylococcus aureus.
CC OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CC OX NCBI_TaxID=158878, 158879, 196620, 1280;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=Mu50 / ATCC 700699, and N315;
CC RC MEDLINE=21311952; PubMed=11418146;
CC RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
CC RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
CC RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
CC RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
CC RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
CC RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
CC RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
CC RT "Whole genome sequencing of methicillin-resistant Staphylococcus
CC RT aureus."
CC RL Lancet 357:1225-1240(2001).
CC [2]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=MW2;
CC RC MEDLINE=22040717; PubMed=12044378;
CC RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
CC RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
CC RA Yamamoto K., Hiramatsu K.;
CC RT "Genome and virulence determinants of high virulence community-
CC RT acquired MRSA."
CC RL Lancet 359:1819-1827(2002).
CC [3]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=NCTC 8325;
CC RC MEDLINE=98241542; PubMed=95731165;
CC RA Wada A., Watanabe H.;
CC RT "Penicillin-binding protein 1 of Staphylococcus aureus is essential
CC RT for growth."
CC RL J. Bacteriol. 180:2759-2765(1998).
CC -!- SIMILARITY: BELONGS TO THE L7AE FAMILY OF RIBOSOMAL PROTEINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003359; BAB56706.1; -.
DR DR
DR EMBL; AP003130; BAB41733.1; -.
DR EMBL; AP004823; BAB94364.1; -.
DR EMBL; U20869; AAC46352.1; ALT_INIT.
DR PIR; B89822; B89822.
DR HAMAP; MF 00574; -; 1.
DR InterPro; IPR000231; Ribosomal_L30e.
DR InterPro; IPR004038; Ribosomal_L7A.
DR InterPro; IPR004037; Ribosomal_L7Ae.
DR Pfam; PF01248; Ribosomal_L7Ae; 1.
DR PROSITE; PS01082; RIBOSOMAL_L7AE; FALSE_NEG.
DR KW Hypothetical protein; Ribosomal protein; Complete proteome.
SQ SEQUENCE 84 AA; 9446 MW; 46EDC013ED029063 CRC64;
CC -----
CC Query Match 60.0%; Score 21; DB 1; Length 84;
CC Best Local Similarity 100.0%; Pred. No. 91;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 8 PVSF 11
CC Db 57 PVSF 60
CC -----
CC RESULT 4
CC Y353 METJA STANDARD; PRT; 84 AA.
CC ID Y353 METJA
CC AC O06918;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein MJ0353.
CC GN MJ0353.
CC OS Methanococcus jannaschii.
CC OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
CC OC Methanocaldococcaceae; Methanocaldococcus.
CC OX NCBI_TaxID=2190;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
CC RC MEDLINE=96337999; PubMed=8688087;
CC RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
CC RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
CC RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
CC RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
CC RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
CC RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
CC RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
CC RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
CC RT jannaschii."
CC RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67488; AAB98342.1; -.
DR DR
DR PIR; A64344; A64344.
DR TIGR; MJ0353; -.
DR KW Hypothetical protein; Complete proteome.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

PM protein - protein search, using sw model

Run on: November 21, 2003, 15:20:39 ; Search time 7.76712 Seconds  
(without alignments)  
598.027 Million cell updates/sec

Title: US-09-010-377-5  
Perfect score: 35  
Sequence: 1 XXXXXXXPVSFXXXXXX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*  
1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	60.0	34	Q96G86	Q96G86 homo sapien
2	21	60.0	35	Q9K016	Q9K016 neisseria m
3	21	60.0	36	Q8EYD2	Q8EYD2 leptospira
4	21	60.0	42	Q8FH25	Q8FH25 escherichia
5	21	60.0	45	P82989	P82989 desulfovibr
6	21	60.0	46	Q8U2K2	Q8U2K2 pyrococcus
7	21	60.0	47	Q9QWE3	Q9QWE3 rattus sp.
8	21	60.0	51	Q9IVY5	Q9IVY5 rotavirus a
9	21	60.0	51	Q9IVY9	Q9IVY9 rotavirus a
10	21	60.0	51	Q9IVY8	Q9IVY8 rotavirus a
11	21	60.0	51	Q9IVZ1	Q9IVZ1 rotavirus a
12	21	60.0	51	Q9IC28	Q9IC28 human rotav
13	21	60.0	51	Q9IVY7	Q9IVY7 rotavirus a
14	21	60.0	51	Q9IVY2	Q9IVY2 rotavirus a
15	21	60.0	51	Q9IVZ0	Q9IVZ0 rotavirus a
16	21	60.0	51	Q9IVY6	Q9IVY6 rotavirus a

17	21	60.0	53	4	Q9P1J0	Q9P1J0 homo sapien
18	21	60.0	54	6	Q8MJA2	Q8MJA2 macaca mula
19	21	60.0	60	2	Q93U58	Q93U58 carsonella
20	21	60.0	61	10	Q8GW73	Q8GW73 arabidopsis
21	21	60.0	61	16	Q8UGC9	Q8UGC9 agrobacteri
22	21	60.0	68	6	O77663	O77663 sus scrofa
23	21	60.0	69	6	Q8HXH4	Q8HXH4 macaca fasc
24	21	60.0	70	9	Q9FZU0	Q9FZU0 pseudomonas
25	21	60.0	71	2	Q8KY51	Q8KY51 bacillus an
26	21	60.0	76	2	Q47103	Q47103 escherichia
27	21	60.0	81	10	Q94D02	Q94D02 oryza sativ
28	21	60.0	81	10	Q8LMD4	Q8LMD4 oryza sativ
29	21	60.0	81	16	Q92RL4	Q92RL4 rhizobium m
30	21	60.0	84	9	Q9G066	Q9G066 bacterioph
31	21	60.0	84	16	Q98R37	Q98R37 rhizobium l
32	21	60.0	84	17	Q8TQ71	Q8TQ71 methanosarc
33	21	60.0	86	10	Q9SKP3	Q9SKP3 arabidopsis
34	21	60.0	86	10	Q8L9D5	Q8L9D5 arabidopsis
35	21	60.0	86	16	Q8CTS9	Q8CTS9 straphylecoc
36	21	60.0	87	2	O52175	O52175 salmonella
37	21	60.0	88	12	Q9JF92	Q9JF92 vaccinia vi
38	21	60.0	89	16	Q98HK4	Q98HK4 rhizobium l
39	21	60.0	91	2	Q9L5F8	Q9L5F8 salmonella
40	21	60.0	92	16	Q935M7	Q935M7 salmonella
41	21	60.0	93	3	Q96TS3	Q96TS3 schizophyll
42	21	60.0	93	16	Q97HM7	Q97HM7 clostridium
43	21	60.0	94	12	Q9DJ76	Q9DJ76 foot-and-mo
44	21	60.0	95	16	Q8FID2	Q8FID2 escherichia
45	21	60.0	96	2	Q9RLG9	Q9RLG9 enterococcu

## ALIGNMENTS

### RESULT 1

Q96G86 PRELIMINARY; PRT; 34 AA.  
AC Q96G86;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC009886; AAH09886.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 34 AA; 3602 MW; 4C7FCC8FE67094B7 CRC64;

Query Match 60.0%; Score 21; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PVSF 11

Db 7 PVSP 10

### RESULT 2

Q9K016 PRELIMINARY; PRT; 35 AA.  
AC Q9K016;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical protein NMS0613.  
GN NMS0613.

```

OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Kachum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citton H., Clark E.B.,
RA Cotton M.D., Uitterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Searlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RT Science 287:1809-1815(2000).
RL EMBL; AE002416; AAF41040.1; -.
RW TIGR; NMB0613; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 35 AA; 4022 MW; A3677246569B6306 CRC64;

Query Match 60.0%; Score 21; DB 16; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 21 PVSF 24

RESULT 3
Q8EYD2 PRELIMINARY; PRT; 36 AA.
ID Q8EYD2
AC Q8EYD2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA4284.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC Ren S.;
ON NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB011582; AAN51482.1; -.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein.
SQ SEQUENCE 36 AA; 4349 MW; 1568528D23C536D5 CRC64;

Query Match 60.0%; Score 21; DB 16; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 21 PVSF 24

RESULT 4
Q8FH25 PRELIMINARY; PRT; 42 AA.
ID Q8FH25
AC Q8FH25;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN C2108.
OS Escherichia coli O6.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AB016761; AAN80568.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 42 AA; 4552 MW; 7A6C730612538776 CRC64;

Query Match 60.0%; Score 21; DB 16; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 27 PVSF 30

RESULT 5
P82989 PRELIMINARY; PRT; 45 AA.
ID P82989
AC P82989;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Cytochrome c3 (Fragment).
OS Desulfovibrio africanus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
ON NCBI_TaxID=873;
RN [1]
RP SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND HEME-BINDING.
RC STRAIN=NCIB 8401 / BENGHAZI;
RX MEDLINE=96162885; PubMed=8573595;
RA Pieulle L., Haladjian J., Bonicel J., Hatchikian E.C.;
RT "Biochemical studies of the c-type cytochromes of the sulfate reducer
RT Desulfovibrio africanus. Characterization of two tetraheme cytochromes
RT c3 with different specificity."
RL Biochim. Biophys. Acta 1273:51-61(1996).
CC -!- FUNCTION: EXCHANGES ELECTRONS WITH [NIFE] HYDROGENASES.
CC PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH PHOSPHORYLATION
CC BY TRANSFERRING ELECTRONS FROM THE ENZYME DEHYDROGENASE TO
CC FERREDOXIN.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
CC -!- PTM: BINDS 8 HEMES.
CC -!- MISCELLANEOUS: ACIDIC TETRAHEME C3 IS NOT A SUBUNIT OF THIS
CC OCTAHEME CYTOCHROME C3.
DR InterPro; IPR000345; CytC_heme_bind.
DR PROSITE; PS00190; CYTOCHROME C; PARTIAL.
KW Electron transport; Sulfate respiration; Heme; Periplasmic.
FT METAL 42 42 IRON (HEME 1 AXIAL LIGAND).
FT METAL 45 45 IRON (HEME 3 AXIAL LIGAND).
FT NON_TER 45 45
SQ SEQUENCE 45 AA; 5114 MW; 1A67207235DD4DB4 CRC64;

Query Match 60.0%; Score 21; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PVSF 11
Db 19 PVSF 22

```